The Sparse Engram of the Lateral Amygdala

by

Dano Morrison

A thesis presented in conformity with the requirements

for the degree of *Master of Science*

Department of Physiology

University of Toronto

The Sparse Engram of the Lateral Amygdala

Dano Morrison

Master of Science

Graduate Department of Physiology

University of Toronto

2016

Memories are stored by discrete changes in the brain, collectively referred to as an engram, that allow activity related to an experience to be reactivated in the future. Associative memories related to threatening stimuli are stored in the lateral amygdala (LA), where the engram is composed of a sparse proportion of neurons. The work in this thesis tests the prediction that this proportion is consistent between memories and investigates mechanisms that may contribute to this consistency. First, it is shown that the proportion of LA engram neurons remains constant between memories of different strengths. Then, it is shown that this proportion increases when inhibitory, parvalbumin-positive interneurons are suppressed during learning. Together, these results provide additional support for the idea memories are stored in sparse assemblies of neurons and suggest a novel role for parvalbumin-positive interneurons in memory formation.

Acknowledgements

I realize that I owe where I am today (completing a MSc. thesis) to a lot of hard work that was done by people who came long before me. They established the scientific ideas and the institutions that provided the fertile ground for my little research project to grow. I simply fussed around in the lab, read a couple papers, and took out some old books from the library to put this thesis together.

It's a modest contribution that I've made, but I hope that I've managed to push the metaphorical cart a little further down the road, ready for the next person to take it on. Specifically, I'd like to thank Dr. Sheena Josselyn for taking a chance with me and inducting me into the ranks of professional scientists with her sage wisdom and savvy mentorship. I also thank Dr. David Ehrlich for taking me under his wing at the start and encouraging my neophytic scientific ideas, Dr. Valentina Mercaldo for her strict and vaguely maternal mentorship throughout my time in the lab, Dr. Adelaide Yiu for being an exemplar of organization, discipline, and dedication (readily apparent from her extremely detailed lab notebooks), Dr. Sung Mo Park for helping me get riled up by the excitement of scientific investigation, Dr. Sherwin Nicholson for making this entire investigation possible with his masterful surgical skills, Dr. Asim Rashid and Chen Yan for their early work that I was able to build upon, Dr. Adam

Table of Contents

List of figures

Chapter 1

Introduction

1.1 The engram

1.1.1 A HISTORY OF IDEAS

If we accept the philosophical (and now scientifically incontrovertible) position that all activity of the mind is mediated by activity of the brain, then it follows that whenever an event is stored in mind such that it can be recalled at a later time (ie. a memory), there must be some concurrent change that occurs in the brain. That change in the brain that allows a memory to be stored is referred to as an engram and has been a topic of interest for neuroscientists for more than a century. Understanding how engrams are formed and how they shape belief, identity, and behaviour is one of the great goals of neuroscience.

Many scientists have devoted their careers to studying this question. Among them, the Canadian neurophysiologist Donald Hebb stands out for his

compelling theory of how information could be encoded into cell assemblies (Hebb, 1949). He was not the first to propose the concept of the engram, the German biologist Richard Semon had done so years earlier (Semon, 1921), but Hebb advanced the field greatly by presenting a theory of memory that drew upon all the neurophysoiological data that had been collected at that time. His thesis, remembered today in the simplified adage *"neurons that fire together, wire together"*, was built on the concept that changes in the strength of connections between neurons lead to the formation of cell assemblies, groups of cells whose coincident activation give rise to certain behaviours or perceptions. Such cell assemblies, formed during the encoding of a memory, increase the chance that the same pattern of activity could be recreated at a later time and trigger memory recall.

Hebb's postulate arrived at a time when many neuroscientists believed that the brain carried out many of its functions holistically, with no particular area of the brain or group of cells being more important than any other. Hebb's own doctoral advisor, Karl Lashley, was perhaps the most distinguished proponent of this idea, coming to this conclusion after his systematic attempt to locate the engram had failed. Lashley had examined the effects of lesions in various cortical areas on the retention of several maze and learning tasks in rodents (Bruce, 2001). He failed to find any specific region of the cortex that, when lesioned, would erase memory. Instead, he only observed memory impairments after large areas of the cortex were damaged, regardless of location. This lead him to propose that memory was broadly distributed throughout 'equipotential' neural circuits and that finding the engram was impossible. This belief would provide a source for debate between the two scientists and within the scientific community until Lashley's death.

It was not until technologies for probing and manipulating the activity of the brain became sufficiently precise that Hebb's ideas could be validated. First, more detailed lesion studies in mammals proved that subcortical areas of the brain were necessary for memory storage (Mishkin, 1978; Morris et al., 1982; Weiskrantz, 1956). Additionally, the first evidence for a discrete molecular change underlying a simple form of learning was discovered in Aplysia, a large sea slug belonging to the phylum *Mollusca* (Kandel, 1978). Most importantly, the discovery of long-term potentiation (LTP) in the hippocampus (Bliss & Lomo, 1973) provided the strongest strongest for Hebb's theory that memories are stored by durable changes in neural connectivity. LTP refers to a phenomenon in which high frequency activity across a synapse increases the strength of that synaptic connection (R. C. Malenka & Bear, 2004). The scientists who discovered LTP wisely left open the question of whether the phenomenon, which was identified in cultured hippocampal cells, could be used by living organisms for memory storage. However, LTP, especially when considered alongside the phenomenon of spike-timing-dependent plasticity, wherein coordinated firing between pre- and post-synaptic neurons drives the induction of plasticity (Markram et al., 1997), corresponds strikingly well with the mechanisms described by Hebb.

1.1.2 OBSERVING THE ENGRAM

More recently, tools that allow neurons active during the time of memory formation to be identified and tagged for future manipulation have been able to answer questions about memory information beyond the level of the synapse. By analyzing the activity of groups of neurons, it has been possible to demon-

strate how neurons activated during memory encoding form neuronal ensembles that underlie memory retrieval. Furthermore, these techniques have allowed an important question at the heart of Lashley and Hebb's debates to be posed: to what degree are engrams distributed throughout the brain?

Strategies for identifying the engram have often relied on immediate early genes (IEGs) such as Fos, Zif268, or Arc. These genes are reliably induced by neural activity (John F. Guzowski et al., 2005) and return to low levels of expression within minutes. Pairing IEGs with expression of molecular tags such as green flourescent protein (GFP) or LacZ allows researchers to permanently label populations of cells active during learning. By combining these permanent tagging approaches with methods to detect IEG expression induced by memory retrieval, researchers have been able to test one key hypothesis made by engram theorists: neurons active during the experience of an event are reactivated during the memory of that event. Several studies have now demonstrated a high degree of overlap between the populations of neurons active during memory encoding and recall (Denny et al., 2014; Reijmers et al., 2007; Tayler et al., 2013). Although this pattern of reactivation could be traced to specific neurons, those neurons were distributed broadly throughout many regions of the brain (Denny et al., 2014; Tayler et al., 2013). Therefore, the engram of a memory appears to involve a widely distributed collection of neurons distinguished on the basis of their activity during the events that lead to the formation of that memory.

1.1.3 RESEARCH SHOWING NECESSITY

Studies that merely observe the engram, however, are unable to demonstrate more than a correlation between the engram and memory. Thus, many studies

have made use of techniques to ablate or silence the neuronal ensembles tagged during memory formation. By eliminating or impairing the ability of the engram to be reactivated, it is possible to demonstrate that engram activity is necessary for memory retrieval.

By selectively introducing inhibitory or toxic receptors into neurons that undergo engram-related changes, many studies have shown that preventing the reactivation of engram cells in specific networks of the brain prevents memory recall. Either through selective genetic ablation (Han et al., 2009), inhibitory optogenetics (Tanaka et al., 2014), or DREADDS (Hsiang et al., 2014), when those neurons that were most active during learning are prevented from firing, memories are either not retrieved or expressed in a weakened state. In comparison to Lashley's lesion experiments, these studies succeeded by specifically targeting subcortical brain regions known to be involved in each particular form of learning under investigation. Furthermore, they were able to target those specific neurons that had presumably undergone changes in synaptic connectivity during encoding.

The ability of localized inhibition to interfere with the reactivation of a widely distributed engram demonstrates the importance of certain regions in initiating memory retrieval. For example, the silencing of only a small subset of cells in the CA1 region of the hippocampus (Tanaka et al., 2014) was sufficient to wipe out a memory that involved distributed representations throughout the brain. These results suggest that distinct areas of the brain, especially those involved in processing contextual or sensory input such as the CA1 or the lateral nucleus of the amygdala, may contain critically important neuronal ensembles that 'trigger' a broad pattern of reactivation across the brain when activated by appropriate retrieval cures. This idea has been supported by the finding

that directly reactivating cortical engram neurons is able to bypass the effects of subcortical inactivation (Cowansage et al., 2014).

1.1.4 RESEARCH SHOWING SUFFICIENCY

In order to demonstrate the sufficiency of engram reactivation for memory retrieval, many researchers have focused their efforts on artificially activating neuronal ensembles formed during learning. In one study, artificially inducing the firing of hippocampal neurons tagged during the formation of a fearful memory was shown to increase the amount of time mice displayed fearful behaviour when placed into a completely novel context (Liu et al., 2012). Many studies have since demonstrated that artificial activation of engram neurons in either the hippocampus or amygdala can trigger of memory recall in the absence of appropriate retrieval cues (Garner et al., 2012; J. Kim et al., 2013a; Yiu et al., 2014). Such artificial memory activation has also been shown to trigger reactivation of widely distributed populations of neurons outside the area of stimulation, such as would be observed during natural memory retrieval (Cowansage et al., 2014).

Studies involving artificial reactivation of memories have been able to go further than simply demonstrating the necessity of engram reactivation for memory retrieval. For example, artificially reactivated memories have been shown to have effects on long-term emotional behaviour (Ramirez et al., 2015). Furthermore, intense artificial activation of engram neurons in the amygdala or hippocampus during exposure to new contexts is capable of modifying engrams and producing false memories (Ohkawa et al., 2015; Redondo et al., 2014). This is a compelling area of research that may contribute greatly to our understand-

ing of how memories interact with ongoing behaviour.

1.1.5 MEMORY ALLOCATION

Understanding how neuronal ensembles are formed during memory formation should provide insight into how information is stored in an engram. Memory allocation refers to the processes by which specific cells and synapses are selected to undergo changes that will become part of the engram. The neural mechanisms by which this occurs have not been specifically defined, but both theoretical and experimental work has proposed that memory allocation is driven by a process of sparse coding that provides efficient information storage.

A critical component of the engram supporting an auditory fear memory can be localized in the lateral nucleus of the amygdala (LA). In the LA, many different sensory input streams converge, allowing associations to be formed between stimuli and the expectation of threat or reward (J. E. LeDoux, 2000). Several experiments exploring auditory fear conditioning in the LA have demonstrated that neurons expressing high levels of the transcription factor CREB are more likely to be allocated to memory ensembles (Han et al., 2007, 2009; Hsiang et al., 2014). CREB is a necessary component of the synaptic plasticity pathways that underlie long-term potentiation (A. J. Silva et al., 1998), but also increases intrinsic neuronal excitability (Y. Zhou et al., 2009). Thus, excitability has been proposed to play an important role in memory allocation, with more excitable cells more likely to be allocated to an engram in the LA. Recent studies have confirmed this hypothesis, showing that manipulating cellular excitability directly without altering CREB expression also increases the likelihood that neurons will become part of an engram (Yiu et al., 2014).

However, artificially increasing the excitability of LA neurons in these studies did not increase the overall size of the LA component of the engram (in terms of number of neurons). This suggests that memory allocation is a competitive process determined by relative rather than absolute excitability (Han et al., 2007; Yiu et al., 2014). In addition, a slew of studies involving different analytical techniques has repeatedly shown that only a small, sparse proportion of principal neurons in the LA become part of any one fear memory trace (An et al., 2012; Ghosh & Chattarji, 2015; Gouty-Colomer et al., 2015; Herry et al., 2008; Quirk et al., 1995; Reijmers et al., 2007; Rumpel et al., 2005) despite the fact that more than 70% of neurons in the LA receive the appropriate sensory innervation (Repa et al., 2001; Romanski et al., 1993). Together, these results suggest that the number of neurons involved in representing sensory events remains constrained to a sparse proportion of neurons despite variations in excitability and sensory input. The concept of small, consistently-sized representations corresponds with theories of sparse coding and data collected from other brain regions (Hromádka et al., 2008; Sanes & Donoghue, 2000; Weliky et al., 2003; Wixted et al., 2014). Sparse distributed coding, in which discrete units of information are encoded across small subsets of neurons in a large network, is thought to provide a structure for high-capacity memory storage that is robust to noise and capable of being implemented in a rapidly changing biological substrate (Ahmad & Hawkins, 2015; Druckmann & Chklovskii, 2012; Krieg & Triesch, 2014). Memory allocation in the LA may proceed in such a way as to promote sparse coding.

1.2   Memory in the brain

Although engram research has been very succesful in describing how memories
are stored at the cellular level, it is important to appreciate that what we think
of as memory in everyday life is actually the product of many different systems.
The brain's many forms of memory can be differentiated on several critera, in-
cluding whether the period of information storage is short or long and whether
information can be consciously retrieved, unconsciously triggered by external
stimuli, or expressed only through the performance of a behaviour. The experi-
mental models used in the engram literature represent only a fraction of these
multiple memory systems, and some types of memories (eg. those that involve
high-level semantic information) are much farther away from being understood
than others. However, there is some empirical and theoretical evidence support-
ing the idea that the fundamental, circuit-level processes of engram formation
are conserved across these multiple memory systems.

1.2.1   WORKING MEMORY

A sharp distinction is drawn between short and long term memories as they
serve different functions are mediated by different forms of processing in the
brain. Short term memory refers to our ability to briefly store and recall, with
a high amount of detail, events that happened recently, on the order of seconds.
The majority of this type of memory has been described as working memory, the
'workspace of consciousness' that allows us to hold on to and manipulate a small
number of items in mind to perform complex cognitive tasks such as reasoning
and language comprehension. As opposed to long-term memory, working mem-
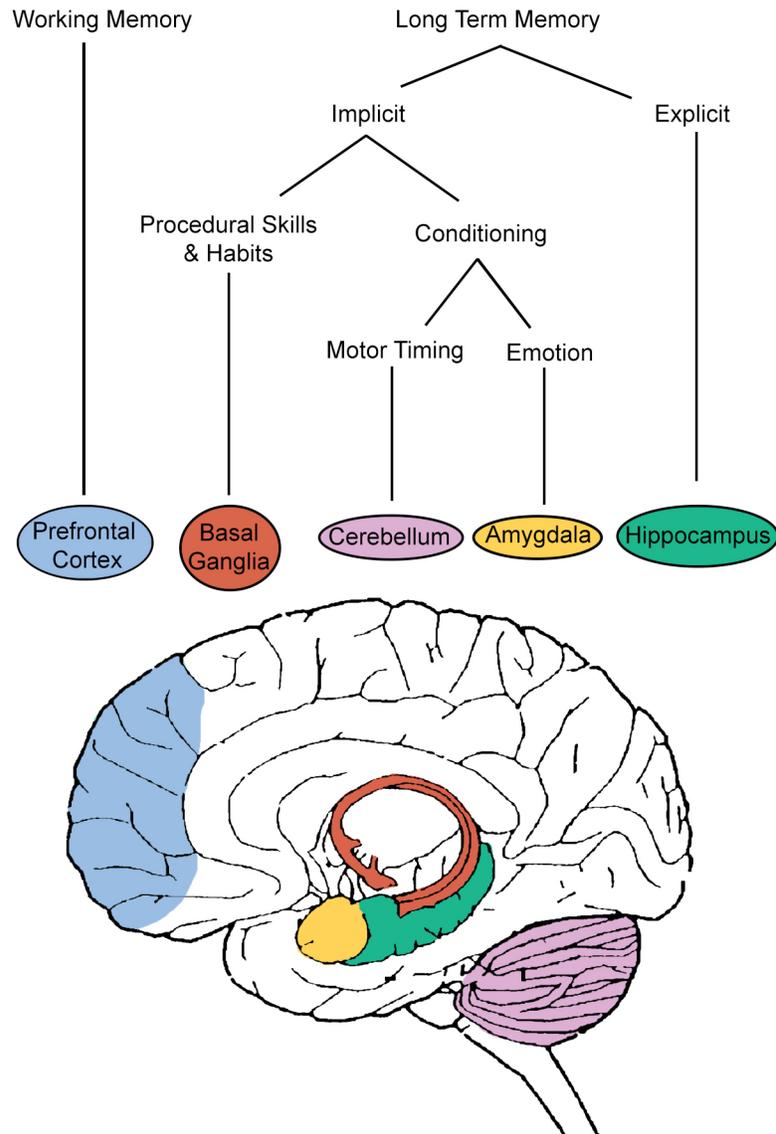
Figure 1.1: **Multiple memory systems in the brain** The most commonly identified forms of memory hierarchically distributed, with the corresponding brain regions supporting those forms of memory indicated. Adapted from Sousa (2016)

ory is dependent on the activity of the cortex, specifically the prefrontal cortex, which is also involved in the processes of attentional control and decision making (Dobbins et al., 2002). Recent research has highlighted the fact that working memory is not solely a prefrontal phenomenon, and that other cortical areas are recruited when particular types of information are maintained in consciousness. For example, keeping the name of someone whom one has recently met in mind involves the continued representation of that name in the temporal areas which process language. However, the prefrontal cortex appears to be critical for selecting, maintaining, and manipulating items in working memory (Postle, 2006). Interestingly, the most influential theories of working memory hold that, due to the flexibility and rapid turnover of items in working memory, information is stored in reverberating loops of neural activiy rather than enduring changes in synaptic connectivity (Wang, 2001). What this entails for the possibility of defining a biophysical engram for working memory is unknown.

1.2.2 Explicit memory

Another large distinction divides long-term memories, those that persist on a timescale of hours to decades. Explicit memories can be explicitly recalled into consciousness at will, while implicit memories rely on external cues or the performance of an action in order to be expressed. These types of memory were shown to be distinct through the study of neurological patients with damage to the medial temporal lobe. These patients, such as the 'Lost Mariner' described in Oliver Sack's *The Man Who Mistook His Wife For A Hat*, demonstrate normal levels of intelligence and can recall events from their childhood, but are entirely unable to form new memories (Sacks, 1998). However, though

they would express no awareness of it, these patients were shown to be capable of certain types of learning, including acquiring complex sensory-motor skills through repeated practice (Cohen & Squire, 1980). This led to the idea that many forms of learning and memory are implicit and are not readily accesible to consciousness or dependent on the medial temporal lobe for formation.

Explicit memories, such as those for events (episodic), facts (semantic), and spatial position, are all critically dependent on the activity of the medial temporal lobe and the hippocampus in particular (Squire et al., 2004). Although many memories undergo a process of consolidation and become independent of hippocampal activity to be recalled (Frankland & Bontempi, 2005), the hippocampus is absolutely crucial for forming explicit memories. The hippocampus does not work alone and draws upon the cortex to recapitulate the sensory and conceptual features of retrieved memories. Thus, it has been proposed that the hippocampus is best thought of as an indexing structure, which does not store memories itself but instead links together networks of cortical neurons that each represent different features of a memory (Teyler & Rudy, 2007). Over time, repeated coactivation of these distributed cortical circuits is thought to link them more tightly together, incorporating memory for events and abstract knowledge into permanent memory stores that can guide future behaviour.

1.2.3 IMPLICIT MEMORY

The ability to acquire skills, habits, and emotional associations falls into the domain of implicit memory. These types of learning are mediated by a number of different systems within the brain, many of them based in subcortical and limbic structures. Importantly, these forms of implicit memory can all be learned and

expressed independent of explicit conscious awareness.

Acquiring procedural skills and habits appears to be dependent on the activity of the basal ganglia, a group of subcortical nuclei that are historically implicated in the initiation of movement and are degenerated during Parkinson's disease. Through repeated practice, the basal ganglia are capable of chunking together sequences of actions, allowing them to be performed more reliably with less conscious effort (Jog et al., 1999). This is related to the concept that we have in everyday life of muscle memory. The chunking of sequences of actions has been shown to reduce the degree of brain activity observed during the performance of complex motor skills, which decreases demand on attentional and motor networks and is an important part of the development of expert-level performance (Debarnot et al., 2014). Importantly, the learned habits that the basal ganglia are involved in acquiring are not restricted just to motor skills. Cognitive skills such as the ability to make probabilistic judgments have also been shown to depend on the basal ganglia and be acquired in a similar way to motor behaviours (Knowlton et al., 1996). Reinforcement learning, in which behaviours that lead to the acquisition of rewards are solidified over time in to habits, is also dependent on the action of the basal ganglia (A. M. Graybiel, 2005).

Classical conditioning, discovered by Ivan Pavlov in his famous experiment with dogs, is another form of implicit learning that is dependent on subcortical structures. The formalized rules of classical conditioning describe how some stimulus that ordinarily has a neutral effect on an organism, the unconditioned stimulus (US), can become associated with a stimulus that produces an innate or instinctual response (CS) with repeated pairing. The responses that can be learned through classical conditioning include certain motor reflexes, for which

learning is acquired and stored by changes in the neural circuits of the cere-bellum (Krupa et al., 1993), and emotional responses that are dependent on plasticity in the amygdala (Janak & Tye, 2015). The amygdala responds to important, intense, and unexpected events by activating regions involved in the expression of emotion and forming associations between otherwise neutral stim-uli and emotional impulses related to the approach of reward or the avoidance of threat. By communicating with other brain regions such as the hippocampus, the amygdala is able to provide emotional context to explicit memories, but its most well-studied role is in the formation of unconscious emotional associations.

1.3   The amygdala

Classical conditioning has formed the basis of much of modern memory research. In animal models, classical conditioning is widely used due to the ease of quanti-fying behaviour with measures such as freezing (an immobile posture indicative of prey behaviour performed to avoid detection by predators). The formation of both fearful and rewarding associations with explicit auditory and visual cues has been shown to depend on the activity of the amygdala, a small subcortical region of the brain that is highly conserved across species (Janak & Tye, 2015).

The amygdala has been shown to be critical for the detection of motiva-tionally salient stimuli in the environment and the expression of emotions such as fear and anxiety. There are two amygdali, situated on either side of the brain in the medial temporal lobe, and although the amygdala is often referred to as a single structure, it can be deconstructed into ~13 distinct subnuclei based on cytoarchitecture and histochemistry. For simplicity, the amygdala is commonly

discussed as consisting of the lateral (LA), basal (BA), and central nucleus (CeA) (P. Sah et al., 2003). These nuclei, collectively referred to as the amygdaloid complex, are separated by small intercalated masses of inhibitory cells (ITCs) that gate the flow of information through the complex, which primarily flows from the LA, through the BA, and either into the CeA, which drives fear and anxiety behaviour related to negative memories, or the nucleus accumbens (NAc), which governs motivation and reward (Ehrlich et al., 2009; S. Lee et al., 2013).

1.3.1 The lateral amygdala

The horn-shaped lateral amygdala is the most dorsally-situated nucleus of the amygdala and is bordered ventrally by the basal nucleus, laterally by the external capsule, and medially by the central nucleus. The LA receives the majority of the sensory projections sent to the amygdala and is thought to function as the amygdala's sensory interface (LeDoux et al., 1990). Somatosensory, gustatory, auditory, and visual cortices all send heavy projections to the LA, as do regions of the thalamus that carry corresponding information. During auditory fear conditioning, auditory information reaches the LA through both associative auditory areas and the medial geniculate nucleus (MGN) of the thalamus. These two input streams are thought to represent two distinct pathways of information: one carrying complex, slow information from the cortex and one carrying early, but simpler signals from the thalamus (Bergstrom & Johnson, 2014). This pattern is similar for visual information, with both cortical and thalamic sources of information synapsing in the LA (C. Shi & Davis, 2001). The LA also receives minor innervation from the prefrontal cortex and hippocampal
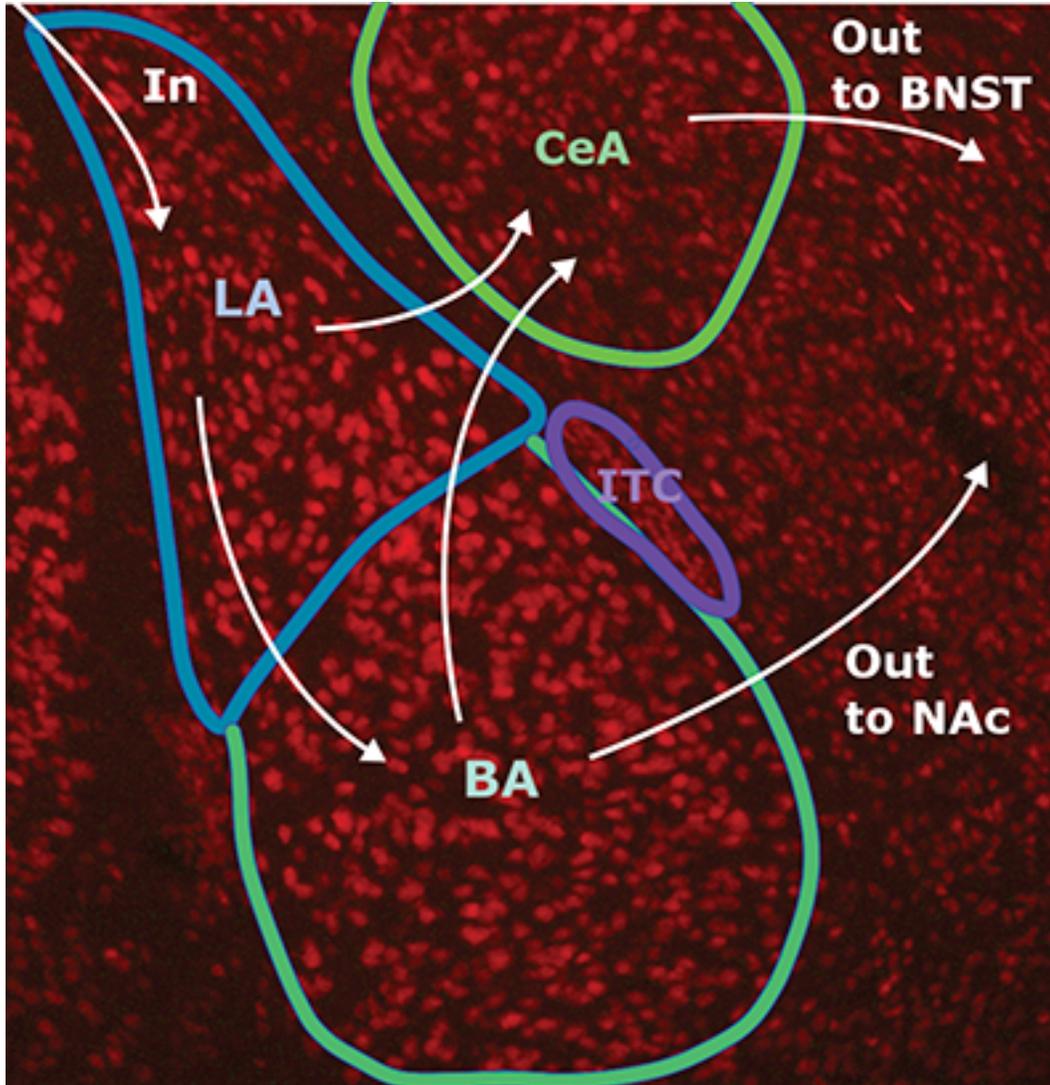
Figure 1.2: **Major nuclei of the amygdala.** The amygdaloid complex with the major subnuclei and patterns of connectivity labeled. During memory recall, sensory input is received by the LA, which sends excitatory projections to both the BA and CeA. The BA, which also receives input from the prefrontal cortex and hippocampus (not shown) projects to the CeA in the case of fearful associations, or to the NAc for rewarding associations. The CeA drives fear and anxiety related behaviour through its connection with the BNST and brain stem. ITC clusters situated between the major subnuclei regulate the flow of information through the circuit

formation (entorhinal cortex and subiculum).

The LA projects to most other regions of the amygdaloid complex, including the basal nucleus, central nucleus, and intercalated cell regions. Interestingly, all of these regions send moderate to heavy input back to the LA, suggesting that most of these connections are reciprocal (Pitkanen, 2000). The purpose that this reciprocal connectivity with the other subnuclei of the amygdala is unknown.

The majority of neurons in the LA (~70%) are cortex-like, spiny pyramidal neurons similar to those found in hippocampus or cortex. However, unlike in those regions, these neurons are randomly organized without any parallel organization in any plane and are naturally quiescent (P. Sah et al., 2003). The remaining neurons primarily consists of a heterogenous population of non-pyramidal local circuit interneurons that provide local inhibition through GABAergic synapses (Spampanato et al., 2011).

1.3.2 SUBNUCLEI OF THE LATERAL AMYGDALA

Although many studies have investigated the nuclei of the amygdala as if they were uniform entities, classic tracing and electrophysiology studies have identified important differences between small subregions, including the LA (P. Sah et al., 2003). The LA has been proposed to consist of three distinct subnuclei: dorsal, ventral-lateral, and ventral-medial. These regions display different patterns of connectivity to targets both within and outside the amygdala. However, how these differences affect information processing during the formation and retrieval of memories is largely unknown and unexplored.

The dorsal subnucleus of the lateral amygdala (LAd), situated at the tip of the LA's horn-like structure receives the majority of the LA's sensory inputs as well as containing the greatest degree of overlap between auditory and somatosensory innervation (Asla Pitkänen et al., 2003; Romanski et al., 1993). The LAd also receives the least amount of reciprocal projections from other nuclei of the amygdala. One study using single-unit recording to measure the activity of single neurons in the LA reported that neurons in the LAd demonstrate the earliest changes in responsiveness to stimuli during fear conditioning (Repa et al., 2001). These early-responding neurons underwent rapid, but transient potentiation during fear learning that decayed as memories became more stable over a number of training events, a finding that was replicated in computational models (Dongbeom Kim et al., 2013a). These observations led the authors to propose that these LAd neurons may be involved in a process of amygdalar memory consolidation dependent on the LAd. However, empirical research since this discovery has failed to replicate or expand upon these findings.

The other two subnuclei of the LA, the ventral-lateral (LAvl) and ventral-medial (LAvm) also receive sensory input, but are distinguished from the LAd by differences in the source of these inputs and in their connectivity to other regions of the amygdala. Generally, the sensory input that the LAvl and LAvm receive is more cortical than thalamic in origin. The LAvm in particular receives most of its input from regions related to higher-order processing, such as the prefrontal and perirhinal cortices, as well as the hippocampus (A. Pitkänen et al., 1997). Furthermore, the LAvl and LAvm both receive local input from the LAd, but do not innervate each other, suggesting that information flows unidirectionally through the LA (A. Pitkänen et al., 1997). These subnuclei

also receive the majority of the reciprocal connections returned to the LA from the BA and CeA.

Despite the differences that have been demonstrated between these subnuclei, it is difficult to draw conclusions as to how they interact and contribute to memory function. For example, some inputs, such as from the dorsal perirhinal cortex, have very little selectivity (C. J. Shi & Cassell, 1999). Furthermore, the high degree of connectivity between these regions suggests that, even if sensory input is initially received by a non-overlapping populations of neurons, it may later be integrated through pathways that link these subnuclei.

1.3.3 THE BASAL AMYGDALA

The basal nucleus of the amygdala is located ventrally to the LA and is thought to integrate input from the LA related to sensory stimuli with information related to context and cognitive processes. Similar to the LA, the BA can be subdivided into distinct subregions: the rostrocellular, magnocellular, and intermediate subnuclei (P. Sah et al., 2003). The BA is highly connected with hippocampal and prefrontal areas (Amano et al., 2011). In addition, the BA contains the majority of the amygdala's projections to the nucleus accumbens that mediate the expression of rewarding memories (Beyeler et al., 2016).

Within the amygdaloid complex, the BA receives many large projections from the LA and sends inputs back to the LA and CeA. The BA also receives projections from thalamic and cortical sensory areas, but the majority of its extra-amygdalar inputs come from the prefrontal cortex and hippocampus, which are thought to contribute to the basal nucleus' role in mediating the extinction and context dependency of learned associations (Herry et al., 2008).

The BA contains a similar variety of neurons as the LA as distinguished by morphology. However, neurons in the BA have been shown to have a significantly higher baseline rate of firing than the LA (P. Sah et al., 2003).

1.3.4 THE CENTRAL AMYGDALA

The central nucleus of the amygdala is located in the dorsomedial region of the amygdaloid complex and consists of two distinct subdivisions, lateral and medial. These subdivisions engage in a great deal of interplay during memory formation and retrieval, with the lateral subdivision receiving the majority of the inputs from the LA and BA and regulating the activity of the medial subdivision (Ehrlich et al., 2009). The CeA is thought to function as the output center of the amygda for fear and anxiety-related behaviour. During retrieval of a fear memory, connections from the medial subdivision of the central amygdala activate the nearby bed nucleus of stria terminalis (BNST), which activates downstream arousal and anxiety-producing regions of the hypothalamus and brain stem (S.-Y. Kim et al., 2013).

The morphological organization of the CeA is unique in comparison to the other amygdaloid nuclei. This is likely due to their different embryological origins, with the CeA sharing its developmental origin with the primarily inhibitory striatum and the LA and BA developing alongside excitatory, cortical regions (P. Sah et al., 2003). Unlike the LA and BA, the CeA is primarily inhibitory and GABAergic, with abundant local connections between subdivisions and a wide variety of interneurons, some of which have been shown to play distinct roles in anxiety (Botta et al., 2015).

1.3.5   Intercalated cell regions

The intercalated cell masses are groups of specialized GABAergic interneurons located in between the lateral and basal nuclei and the central nucleus. Although not much is known about these intercalated cells, they have been shown to provide both feed-forward and feed-back inhibition to the other nuclei of the amygdala and are thought to function as an inhibitory gate that regulates the flow of information between the input and output interfaces of the amygdala (Pape, 2005).

1.4   Inhibitory Interneurons

The amygdala, much like any other brain region, is highly dependent on an appropriate balance between inhibitory and excitatory neural activity. Throughout the amygdaloid complex, there is a wide variety of inhibitory interneurons with various physiological properties and patterns of connectivity (Ehrlich et al., 2009; P. Sah et al., 2003; Spampanato et al., 2011; Waclaw et al., 2010). In the LA and BA, these interneurons comprise about 20% of the neuronal population and have been shown to be central to the process of fear conditioning (Ehrlich et al., 2009), coordinating the flow of information between nuclei and gating sensory input through finely-tuned systems of inhibition and disinhibition (Wolff et al., 2014). These interneurons can be distinguished on the basis of the presence of different molecular markers, including parvalbumin (PV), somatostatin (SOM) and cholecystokinin (CCK), which are strongly correlated with different functional characteristics.

1.4.1 Parvalbumin-containing interneurons

Interneurons that express the calcium-binding protein parvalbumin are prevalent (~15% of neurons) in the lateral and basal nucleus of both the rodent and human amygdala (Sorvari et al., 1996). PV^+ cells display a heterogeneity of firing properties (D. G. Rainnie et al., 2006), but generally share the same morphological characteristics. PV^+ interneurons almost always form synapses on the cell bodies of principal neurons, where they are able to effectively modulate neural excitability and mediate synchronous oscillations (Freund & Katona, 2007). PV^+ interneurons primarily receive input from local excitatory neurons. However, they have been shown to be highly sensitive to sensory input from both thalamus and cortex, suggesting that they may receive cortical or thalamic innervation as well (Szinyei et al., 2000; Woodson et al., 2000).

1.4.2 Somatostatin-containing interneurons

Somatostatin-containing interneurons represent a similar proportion of the interneuronal population as those expressing PV (Sosulina et al., 2010). SOM^+ interneurons display a distinct delayed, stuttering pattern of firing in response to stimulation. Importantly, SOM^+ neurons form synapses on the distal dendrites of pyramidal neurons rather than on the cell body, which lends them the ability to selectively inhibit input arriving on individual dendritic branches rather than manipulate whole-cell excitability.

As the dominant interneuronal subtypes in the amygdala and hippocampus, SOM^+ and PV^+ interneurons interact with each other to a large extent. It has been suggested that PV^+ interneuron-mediated inhibition of SOM^+ cells

gives rise to a process of disinhibitory sensory control that is involved in the formation of associative memories in both the hippocampus (Lovett-Barron et al., 2012) and LA (Wolff et al., 2014). In both these regions, early activation of PV^+ interneurons has been proposed to inhibit downstream SOM^+ interneurons, thereby enhancing responsiveness to sensory information delivered to the distal dendrites of pyramidal neurons and promoting the excitability necessary for plasticity and allocation to the engram.

1.4.3 CHOLECYSTOKININ-CONTAINING INTERNEURONS

Interneurons in the LA that express cholecystokinin are perhaps the most diverse in terms of their functional properties (Jasnow et al., 2009). For example, they have been shown to demonstrate both adapting and non-adapting firing, high and low input resistance, as well as both fast and slow after-hyperpolarization periods. This variety may be related to the fact that mRNA for CCK is more broadly distributed throughout the region than any other molecular marker used to distinguish interneuronal subtypes. In the amygdala, not much is known about the functional role of CCK^+ interneurons. However, there has been some suggestion that a distinct subtype of CCK^+ cells may be specifically involved in mediating anxiety related behaviour (Truitt et al., 2009).

1.5 Modeling memory

In order to understand a system as complex as the brain, it is necessary to develop models that present simplified versions of phenomena. By reducing the amount of variables in a model to those that are specifically relevant to the

function or task at hand, it becomes much easier to examine the relationship between relevant factors and test the predictive or explanatory value of theories about how a system works. Such models can be especially useful especially useful when they create novel predictions that can then be tested experimentally. Since the time of Donald Hebb, theoretical models in the field of learning and memory have proven their usefulness for understanding how changes in the brain give rise to memory.

Most computational models used to study learning and memory belong to a class of models known as 'neural networks'. These networks are composed nodes, mathematical abstractions of neurons that are connected to each other through connections (often referred to as edges) that represent simplified synapses. Each node in a neural network has an activation level that is determined by the input into that node from outside the network, or from other nodes in the network. A node's activation level can be thought of as the likelihood that it will generate an output of its own or, alternatively, how active a neuron is on a scale from 0 (not active) to 10 (fully active). The edges in the neural network transmit this activation level between nodes, with the amount of activity that is sent from the output node to the input node determined by that edge's synaptic weight. Importantly, weights can have both positive and negative values, corresponding to excitatory and inhibitory connections in the brain (although some models allow nodes to have output edges with both positive and negative weights, which is biologically inaccurate). Each set of nodes that sends output to another set of nodes is referred to as a layer. The result of this simple architecture is that each node is capable of computing a weighted sum from a number of inputs. Although this low-level operation is very limited on its own, as more nodes (and layers of nodes) are added to a neural network,

and as rules are implemented that allow the connections between nodes to be modified, their computational power becomes increasingly powerful. In addition to being able to solve computationally 'hard' problems such as mastering the game of Go (Silver et al., 2016), neural networks have been used to successfully model many forms of learning and memory, from Pavlovian conditioning (M. S. Fanselow, 1998) to systems-level consolidation of long-term episodic memories (Santoro, 2015).

1.5.1 Autoassociative neural networks

Hebb's original prediction for how memories could be stored in the brain involved the simple rule that if cell A connects to a second cell B, and if the two cells are repeatedly active at the same time, then the connection between them is strengthened, so future activity in A is more likely to lead to activation in B (Hebb, 1949). This rule, now known as Hebbian learning can be clearly replicated in neural networks in which the strength of connections between nodes is increased if they are active at the same time.

One class of neural network that implements Hebbian learning are autoassociative netoworks. The distinguishing characteristic of these networks is that each node is both an input and an output node, meaning that it both receives input from external sources and contributes to the overall output of the nework by transmitting its activity to other nodes (Gluck & Myers, 2001). When a pattern of input activates an autoassociative network, some subset of nodes are activated by that external input, while others remain silent. Which neurons are activated is determined both by the connectivity and weights of external inputs and, because nodes are able to excite each other, pre-existing

weights between nodes in the network (Figure 1.3a). If a novel pattern of external input is received, Hebbian learning in the autoassociative network ensures that the synaptic weights of edges between coactive nodes are strengthened such that the subset of neurons activated by external input are bound together into an ensemble, storing this pattern in the network (Figure 1.3b,c). Later, if an incomplete or distorted version of the same input is given to the network, this pattern is more likely to be reinstated due to the enhanced connectivity between those neurons that were most excited by the original input. This process, known as pattern completion, is thought to play an important role in the brain's ability to recognize patterns in our environment and recall previous sensory representations during classical conditioning (ie. if a tone and shock are presented simultaneously, nodes that respond selectively to either stimulus will be linked such that later presentation of either tone or shock alone will lead reactivation to the entire population that was active when both tone and shock were presented).

1.5.2 Sparse distributed memory

The theory of sparse distributed memory is a mathematical model that attempts to describe how memory could be implemented in the brain with a formal, quantified approach. Sparse distributed memory does not require neural networks (indeed, it has been widely implemented in the world of traditional computer science), but it is very easily implemented by neural networks, including autoassociative networks. The central principle of sparse distributed memory is the statistical recall of previously active patterns distributed across a range of different locations in a very large network (Kanerva, 1988).
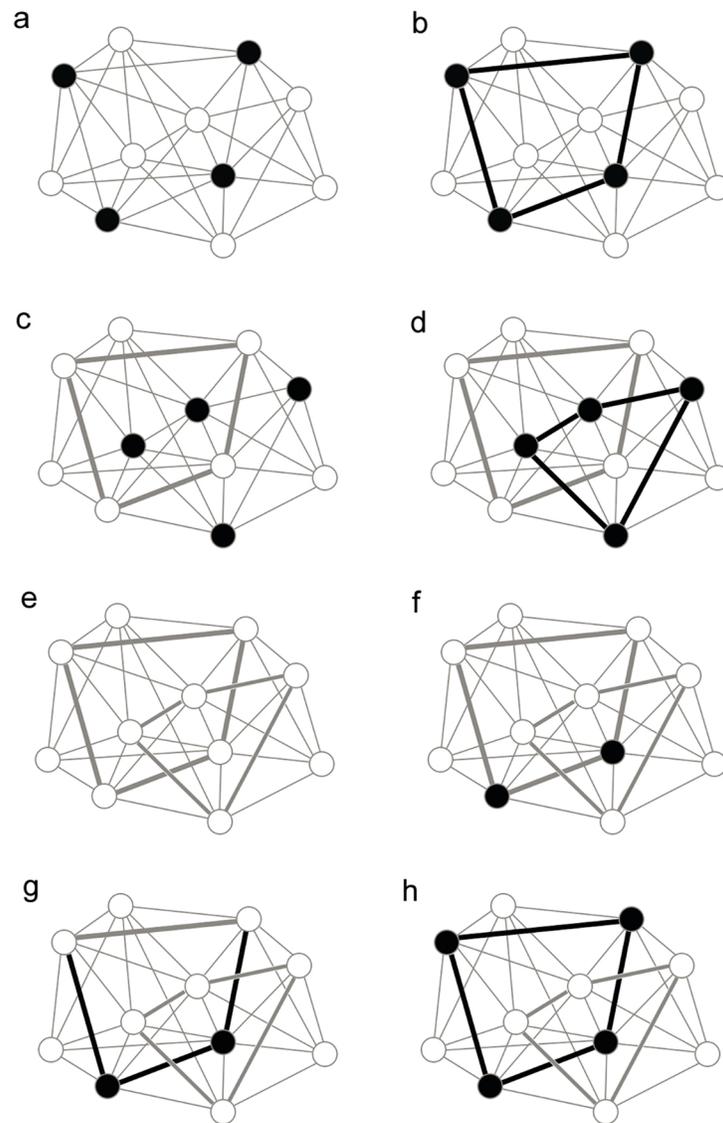
Figure 1.3: **Autoassociative memory.** Illustration of autoassociative memory encoding and retrieval. (a) Sensory input (not shown) to the network activates a certain subset of nodes (solid circles). (b) Hebbian learning leads to the strengthening of the weightened connections between nodes that are activated by sensory input, storing this pattern of activity. (c) Later, a different new pattern of sensory stimuli evokes the activation of a different subset of nodes in the network. (d) Weights between nodes active during this new pattern are strengthened and (e) both patterns are now stored in the network (heavy lines). (f) If a partial version of a previously stored pattern is presented to the netowork, some of the nodes involved in that pattern will be activated. (g) Activation will then flow along previously strengthened connections and lead to (h) the reactivation of the previously stored pattern. Adapted from Gluck and Myers, 2001.

In order to understand sparse distributed representations, it is best to draw comparisons with other coding schemes such as local coding. In a local coding, every unit of information is stored at a single location with no overlap between units. A common analogy to this is a computer keyboard, where each key matches to a character in a one-to-one fashion. In neuroscience, local coding can be expressed to the extreme by the "Brad Pitt neuron" (or "grandmother cell") hypothesis. In this hypothesis, there exist neurons that store memories for only one particular person, place, or thing. When, for example, an image of Brad Pitt is shown, the recognition of Brad Pitt is mediated entirely by the firing of a single "Brad Pitt neuron". Local representations have the advantage of being easy to decode (the memory is always in one location) and they avoid the risk of interference between memories. However, they have extremely low capacity, as every memory requires its own dedicated location, and they are inflexible, as new memory addresses would need to be added to permit new learning. Furthermore, because memories are disconnected from each other, such representations do not allow memories to be linked or generalized in any way. This lack of flexibility would be completely untenable for an organism required to navigate a noisy world of fuzzy inputs.

Another example of a coding scheme that is inefficient within in the hardware of the brain is dense coding. This is the form of memory implemented by computer binary that stores information as patterns distributed across the entirety of a fixed number of locations. For example, in ASCII code, each letter is encoded as a string of 8 binary values. When this string is read, the computer then reads each bit and performs a lookup operation to determine which character corresponds to the string that has been read (01100001 = a, 01100010 =b, etc.). Such a coding scheme is convenient for use in computers because only

a small number of bits need to be read to encode a large number of distinct memories (8 bits = 256 possible combinations). However, there are significant drawbacks to this scheme outside of the orderly, immutable world of computer memory. First, because all possible locations (bits) must be read in order to perform the lookup operation, only one pattern can be present in the network at one time. Furthermore, dense representations are very sensitive to noise. Each pattern of active locations represents a distinct memory, so if the activity at one location is altered, the memory that is retrieved will be completely changed.

The substrate of the brain is very different from a computer: cells and synapses are constantly being modified or turned over, and the input cues that networks receive during memory retrieval are noisy. In this environment, the most effective coding scheme is sparse distributed memory, where information is spread across a sparse subset of units in a larger total population. At their most basic form, sparse distributed memory networks can be represented as very large sets of locations, with only a small percentage of those locations active during the representation of any one piece of information. Because there is so much 'room' in the set of locations, sparse codes can coexist in the same network. Furthermore, each individual location can be involved in the representation of multiple memories. Overlap between patterns can be used to encode associations between memories. Additionally, memories can be recalled from incomplete or corrupted input patterns since presenting the network with partial pattern is able to narrow the range of appropriate stored patterns to practical degrees of statistical certainty (Ahmad & Hawkins, 2015). Sparse distributed representations also allow new patterns to be stored simply by altering the connectivity of existing units, such as in a neural network. No new addresses need to be created to store new patterns. Rather, the network merely needs to be

modified enough to create a new pattern of activity that is sufficiently different from pre-existing patterns. These features, along with neurophysiological data showing sparse patterns of activity during memory retrieval in the amygdala (Bach et al., 2011) or hippocampus (Wixted et al., 2014) provide strong evidence for sparse coding in the brain.

1.5.3 Modeling memory allocation in the lateral amygdala

In order to understand the mechanisms contributing to the induction and storage of emotional memories, a significant amount of work has been done to develop neural network models of the LA. One such model, introduced by D. Kim, Paré, & Nair, consists of an autoassociate neural network that contains both excitatory and inhibitory cells, implements Hebbian forms of long-term potentiation and depression, and replicates patterns of intra-LA connectivity discovered by systematic electrophysiological recording studies in the rat LA. By delivering input to the network corresponding to what would occur during auditory fear conditioning, the group has been able to replicate the findings of several electrophysiological and IEG-based studies and has gone on to make proposals about the mechanisms by which neurons are allocated to the engram during encoding.

A series of experiments in mice has demonstrated that neuronal excitability plays a central role in biasing allocation to an engram (Han et al., 2007; Hsiang et al., 2014; Yiu et al., 2014; Y. Zhou et al., 2009). However, artificially increased excitability of up to 20% of LA neurons in these studies did not increase the overall size of the of the LA component of the engram (in

terms of number of neurons), suggesting that allocation is a competitive process. However, these studies were unable to explain how this process of competition occurs or what purpose it serves. In order to answer these questions, Kim et al. performed a series of experiments in their model, arbitrarily increasing the excitability of subsets of excitatory neurons and recording the effects on the size of the stable engram after repeated training in an auditory fear conditioning paradigm. Interestingly, they were able to replicate and extend the effect demonstrated in previous studies: increasing excitability in up to 25% of cells did not alter the size of the memory trace (Dongbeom Kim et al., 2013c). Furthermore, with the control their model afforded them, they were able to demonstrate that inhibition is the primary factor in constraining neuronal allocation to an engram (Feng et al., 2016). Their model demonstrates that competitive interactions between principal neurons are mediated by disynaptic connections involving inhibitory interneurons. Essentially, highly excitable principal neurons activated during learning inhibit many of their principal neuron neighbors via indirect, inhibitory connections. Highly excitable principal neurons that share excitatory connections with each other become even more excited, while principal neurons that primarily receive disynaptic inhibitory connections from this highly excitable subset are silenced. This process magnifies the difference in excitability between 'winner' and 'loser' neurons. Furthermore, they demonstrate that plasticity at inhibitory synapses is able to compensate for increases in excitability such as would be seen during CREB overexpression.

This neural network modelling approach was also able to suggest why engrams in the LA are constrained. First, the sparse code enforced by competitive interactions is consistent with theoretical models of sparse distributed coding schemes that have also been demonstrated in other areas of the brain, such as

hippocampus (Wixted et al., 2014) and cortex (Gdalyahu et al., 2012) Thus, it reasonable to assume that the LA benefits from the many advantages of sparse distributed representations in a neural substrate (high capacity, robustness to noise). Second, the model revealed that sparsity enforced by competitive interactions could ensure an optimal degree of stimulus specificity (D. Kim et al., 2015). By selectively modulating plasticity of either excitatory or inhibitory synapses, the researchers demonstrated that excitatory connectivity drives the number of cells allocated to the engram upwards, but increases the probability that memories will be generalized. Inhibitory connectivity had the opposite effect, decreasing the proportion of engram cells and making memories more specific. For an animal in the wild, the balance between memory specificity and generalizability is extremely important, drawing the line between failure to respond to threats and chronic anxiety.

Chapter 2

Hypothesis

The goal of this thesis is two part. First, to test whether the proportion of lateral amygdala neurons allocated to an engram remains constant despite variations in memory strength. Second, to examine whether inhibitory interneurons are implicated in the mechanisms that ensure this sparse coding. To test these hypotheses, I will manipulate memory strength in an auditory fear conditioning paradigm and measure the proportion of neurons in the LA allocated to an engram of conditioned fear. I will then perform a similar experiment including the inactivation a subset of inhibitory PV^+ interneurons in the LA during fear conditioning. If these inhibitory cells are involved in a mechanism that constrains the engram, I predict that this manipulation will alter the number of the neurons allocated to the engram. Together, these results should provide novel evidence for the existence of memory allocation mechanisms in the LA that involve inhibitory interneurons.

Chapter 3

Methods

3.1   Mice

Except where noted, adult (at least 8 weeks of age) male and female F1 hybrid (C57BL/6NTac X 129S6/SvEvTac) wild-type (WT) mice were used in all experiments. Mice were bred at the Hospital for Sick Children, provided with food and water ad libitum, and group housed (4 per cage) on a 12 h light/dark cycle. All mice were briefly single housed, transported, and handled once daily for 6 days prior to experiments. All experiments were performed during light-phase. All procedures were conducted in accordance with policies of the Hospital for Sick Children Animal Care and Use Committee and conformed to both Canadian Council on Animal Care (CCAC) and National Institutes of Health (NIH) Guidelines on Care and Use of Laboratory Animals. For experiments involving PV^+ interneuron inhibition, heterozygous PV-Cre knockin driver mice (B6;129P2-Pvalbtm1(cre)Arbr/J) maintained on a C57BL/6 genetic background were used. These mice express Cre recombinase in PV neurons, without disrupting endoge-

nous PV expression. PV-Cre mice were obtained from Jackson Labs and were originally generated by Silvia Arber, FMI (Hippenmeyer et al., 2005).

3.2 Auditory fear conditioning

Auditory fear training involved placing mice into a conditioning chamber and, 2 minutes later, presenting a tone (2,800 Hz, 85 dB, 30 s) that co-terminated with a 2s footshock. Testing was performed 24 hours later by placing mice in the fear chamber with novel contextual cues and, after a 2 minute baseline period, presenting an identical tone for 1 minute. During testing, an overhead camera captured the animal's activity and automated motion analysis was used to assess percentage time spent freezing during the tone (defined as an immobilized, crouched position with an absence of any movement except respiration (M. S. Fanselow & Bolles, 1979).

3.3 Immunohistochemistry

We used the expression of Arc protein in neurons active following training or a memory test as a proxy for identifying neurons that compose part of an engram supporting that memory. Arc protein is reliably elevated in excitatory neurons 90 min after activity (John F. Guzowski et al., 2005) and has been previously used as a proxy for visualizing neuronal components of an engram (Gouty-Colomer et al., 2015). 90 minutes after the delivery of the auditory tone CS during training or testing, mice were transcardially perfused with 4% paraformaldehyde (PFA) and brains removed. Brains remained in PFA for 1

day and were then transferred into cryoprotectant solution (20% sucrose) at 4 ºC. Three days later, brains were sectioned into 50 m slices using a cryostat with an inter-slice interval of 100 um (taking every third slice). Slices were then washed and incubated overnight with Arc polyclonal rabbit antibody (1:500, Synaptic Systems, Cat. No. 156 003) at 4 ºC. In the experiment involving manipulations of shock intensity, slices were also incubated with NeuN polyclonal mouse antibody (1:1000, Millipore, MAB377). Staining was visualized either with fluorescent secondary goat anti-rabbit Alexa 288 (1:1000, Life Technologies, Cat. No. A11034) and goat anti-mouse Alexa 568 (1:1000, Life Technologies, Cat. No. A11004) antibodies or 3,3 -Diaminobenzidine (DAB) or in the case of the first experiment. Measures of Arc$^+$ cell density by area in the first experiment using DAB were obtained by imaging sections under a 10x Nikon light microscope and manually counting the number of Arc$^+$ cells in these images with ImageJ software (NIH) (Schneider et al., 2012) by experimenters blind to the treatment condition. ANOVAs with post-hoc Fisher's LSD tests were conducted to compare Arc$^+$ cell density between different treatment groups within each amygdala nucleus.

3.4   AAV viral vectors

The AAV8-hSyn-DIO-hM4Di-mCherry vector was obtained from the UNC Vector Core (Chapel Hill, NC). This adeno-associated virus (AAV) contains a double floxed inverse open reading frame of hM4Di fused to mCherry that is expressed from the human synapsin (hSyn) promoter after recombination by Cre recombinase (Krashes et al., 2011). The average titer of the virus stocks was 4.0 $\times$ 107 infectious units/ml. Mice were allowed to recover for 4 weeks for maximal

viral expression before performing behavioural experiments.

3.5 Surgery

Before surgery, mice were pre-treated with atropine sulfate (0.1 mg/kg, ip), anesthetized with chloral hydrate (400 mg/kg, ip) and placed in a stereotaxic frame. During surgery, cranial skin was retracted and holes were drilled in the skull above the amygdala (AP = -1.4 mm, ML = ±3.4 mm, V = -5.0 mm from bregma). AAV-DIO-hM4Di-mCherry (1.0 l) was infused bilaterally into the amygdala. After surgery, mice recovered for 4 hours before being returned to their standard housing.

3.6 Cellular compartment analysis of temporal activity by fluorescent in situ hybridization (catFISH)

Mice were trained with various shock conditions (1 x 0.4 mA, 1 x 0.7 mA, and 3 x 0.7 mA) in order to create memories of different strengths. A group that received a shock immediately upon entering the chamber was included as a control. 5 minutes after either training or testing, animals were sacrificed and brains were subsequently processed with catFISH. Brains were then imaged and analyzed with systematic stereological procedures to identify the proportion of arc+ neurons in the LA. We performed catFISH for *Arc* mRNA to identify neurons activated during fear conditioning or retrieval. 5 min after either auditory fear conditioning or testing, brains were removed and frozen. Tissue was then sectioned (20 µm) and prepared for FISH according to previously described pro-

tocols (Han et al., 2007). Visualization of *Arc* mRNA was performed through the use of a DIG-conjugated anti-sense probe corresponding to the *Arc* open reading frame. In order to minimize differences in staining, sections from 4 mice were mounted together on one slide. Following hybridization and amplification of *Arc* and H1a signals, counterstaining with Hoechst 33258 to visualize nuclei was performed.

Following catFISH processing, sections were imaged on a laser confocal microscope (Zeiss LSM 710) to obtain optical z-stack series with a step size 1 µm apart. Stacks were analyzed for nuclear *Arc* by two individuals unaware of treatment condition. Stereological counting was then performed by two individuals unaware of treatment condition, with at least four sections counted for each mouse. ANOVAs with post-hoc Tukey's tests were conducted to compare *Arc+* cell density between different treatment groups.

3.7   Stereological counting

To accurately determine the proportion of neurons active in the LA during auditory fear training and testing, we used unbiased stereological principles and systematic sampling techniques to obtain the ratio of Arc-expressing neurons to the number of total neurons. For the experiment involving manipulation of shock intensity, we used expression of the neuron-specific protein NeuN to obtain a measure of the neuronal population directly. However, for experiments involving catFISH and AAV-DIO-hM4Di-mCherry, the nuclear stain 4',6-diamidino-2-phenylindole (DAPI) was used to obtain a measure of the total cellular population that was later adjusted based on estimates of the proportion of cells in the

LA that are neurons. Several sections containing both DAPI and NeuN were analyzed to obtain this empirical estimate (78.9% ± 2.3, n = 10). Stereological counts were adjusted by this ratio without affecting the relationships between groups.

Counting was performed by an experimenter unaware of the treatment condition using the Optical Fractionator probe within Stereo Investigator (version 10, MBF Bioscience, Williston, VT USA) on every third section (50 m thickness, at least 5 sections per mouse) from the left hemisphere. A counting frame of 120 m x 120 m was randomly distributed according to a 250 m x 250 m grid throughout the LA. Section thickness was recorded at every sampling site to compensate for any potential thickness variation across the sections due to tissue processing. A 12 m dissector height was used with 2 m guard zones placed at the top and bottom of each section. Coefficient of Error values, determined by Gundersen's method (Gundersen and Jensen, 1987) (m=1) were < 0.11 in all samples. For analysis of LA subnuclei, contours were drawn between regions as described by the stereotaxic mouse brain atlas by Paxinos & Franklin (2001). ANOVAs with post-hoc Tukey's tests were performed to compare Arc+ cell density between treatment. All stereological data was collected only from the left hemisphere in order to avoid potential hemispheric differences in Arc activity, although preliminary results showed no significant difference in Arc+ cell density between hemispheres (paired t-test, P > 0.1).

3.8   Clozapine-N-oxide (CNO)

CNO (Toronto Research Chemicals [TRC]) was prepared in a stock solution of 10 mg/ml in DMSO and subsequently diluted to a working concentration of 1 mg/ml in phosphate-buffered saline (PBS). CNO was injected intraperitoneally at a dosage of 2.0 mg/kg 1hr before training. Vehicle solution used as a control consisted solely of PBS.

Chapter 4

Results

4.1 Arc expression increases in the amygdala in response to encoding and retrieval of conditioned fear

Similar to previous studies (Gouty-Colomer et al., 2015; Reijmers et al., 2007; Tayler et al., 2013), Arc protein expression was visualized 90 min after memory retrieval to identify neurons that were active during memory encoding or retrieval as part of an engram. Specifically, mice were trained in auditory fear conditioning and the number of neurons positive for Arc (Arc^+) 90 min following various manipulations (e.g., initial training or memory recall) was examined (Figure 1a). Arc^+ neuron density (per 100 μm^2) was assessed in the LA, BA, central amygdala (CeA) and intercalated cell masses (ITC) (Figure 4.1b). Non-learning control groups were also included (homecage mice, mice exposed to tone alone, chamber alone or an immediate shock) (Frankland et al., 2004; Zelikowsky et al., 2014).

There was a significant increase in the number of Arc^+ neurons throughout the amygdala after both training and testing (Figure 4.1c). Arc^+ cell density was elevated both after training and testing (in comparison to homecage controls) in the LA [$F(5, 36) = 10.63$, $P < 0.001$], BA [$F(5, 36) = 5.20$, $P < 0.001$] and CeA [$F(5, 36) = 4.99$, $P < 0.001$]. However, there was no significant difference in Arc^+ density between treatments in the ITC region [$F(5, 36) = 1.86$, $P > 0.05$]. Importantly, the control tone alone, chamber alone or immediate shock groups did not differ from the homecage group in any amygdala region ($P > 0.1$). Fear training and testing induced a similar number of Arc^+ neurons in many amygdala regions (LA, $P > 0.5$; BA, $P > 0.1$; CeA, $P > 0.5$; ITC, $P > 0.5$), suggesting that a similar number of neurons were active during encoding and retrieval of an auditory fear memory.

4.2 Memory retrieval activates a constant proportion of cells in the amygdala despite varying memory strength

In order to assess the relationship between memory strength and the size of the Arc^+ engram, the intensity of the CS footshock delivered during auditory fear conditioning was manipulated (0.3 mA, 0.5 mA, 0.75 mA). Following a memory test delivered 24 hr after fear conditioning, the percentage of neurons involved in the engram was assessed through immunohistochemical staining for Arc protein and the neuron-specific NeuN protein (Duan et al., 2016) with systematic stereological counting performed to quantify the ratio between these two labeled populations.

There was a consistent proportion of Arc^+ neurons (10-15%) in the LA
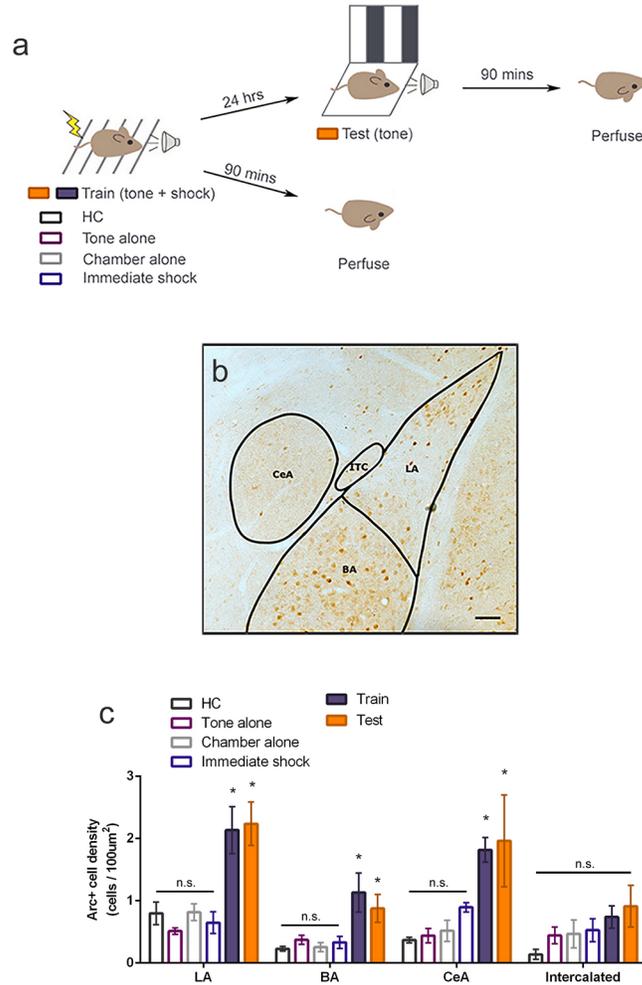
Figure 4.1: **Increase in Arc expression throughout the amygdala after auditory fear training and testing.** (a) Experimental procedure. Mice in the Train group were perfused 90 mins after fear condition in which a tone was paired with a footshock. Tone Alone, Chamber Alone and Immediate shock groups were similarly treated and perfused 90 min later . The test group was perfused 90 mins after a testing session in a novel context 24 hr after initial training. (b) Representative image of the amygdala after immunohistochemical Arc staining with amygdaloid nuclei defined according to Paxinos and Franklin. Scale bar = 50 m. (c) Arc+ cell density increased following training and testing groups in all nuclei of the amygdala except the intercalated regions (n = 8, HC; n = 8, Tone alone; n = 7, Chamber alone; n = 5, Immediate shock; n = 7, Train; n = 7, Test). * P < 0.05, ** P < 0.001 for Fisher's LSD post-hoc test for multiple comparisons. Data is expressed as mean ± sem.

regardless of memory strength. Although varying shock intensity produced different levels of freezing during the test (Figure 4.2a; $F(2, 14) = 4.39$, $P > 0.05$) consistent with different strengths of memory, no difference in the proportion of Arc^+ neurons was observed between groups that received fear conditioning (Figure 4.2b; $F(2, 14) = 0.22$, $P > 0.5$). Arc^+ proportion increased in comparison to homecage controls in all trained groups (Tukey's post-hoc). Furthermore, linear regression analysis showed no relationship between the amount time a mouse spent freezing to the tone and the proportion of Arc^+ cells in the LA (Figure 4.2c, $R2 = .003$). Together, these data suggest that a similar number of neurons are allocated to the engram regardless of the intensity of training conditions or the strength of learned fear associations.

4.3 *Arc* mRNA increases in a stable proportion of LA neurons during fear conditioning

In order to gain a more temporally precise indication of which neurons in the LA were reactivated as part of on engram, we measured Arc mRNA with cellular compartmental analysis of temporal activity by fluorescence in situ hybridization (catFISH). Localization of Arc mRNA in the nucleus is a molecular signature of a neuron that was active in the previous 5 min (John F. Guzowski et al., 2005). Therefore, visualization of Arc mRNA allowed us to identify neurons active during memory retrieval with a high degree of temporal accuracy.

Similar to the previous experiment, mice were fear conditioned with different training intensities (1 x 0.4 mA, 1 x 0.7 mA and 3 x 0.7 mA footshock). As expected, these training conditions produced varying degrees of freezing to the
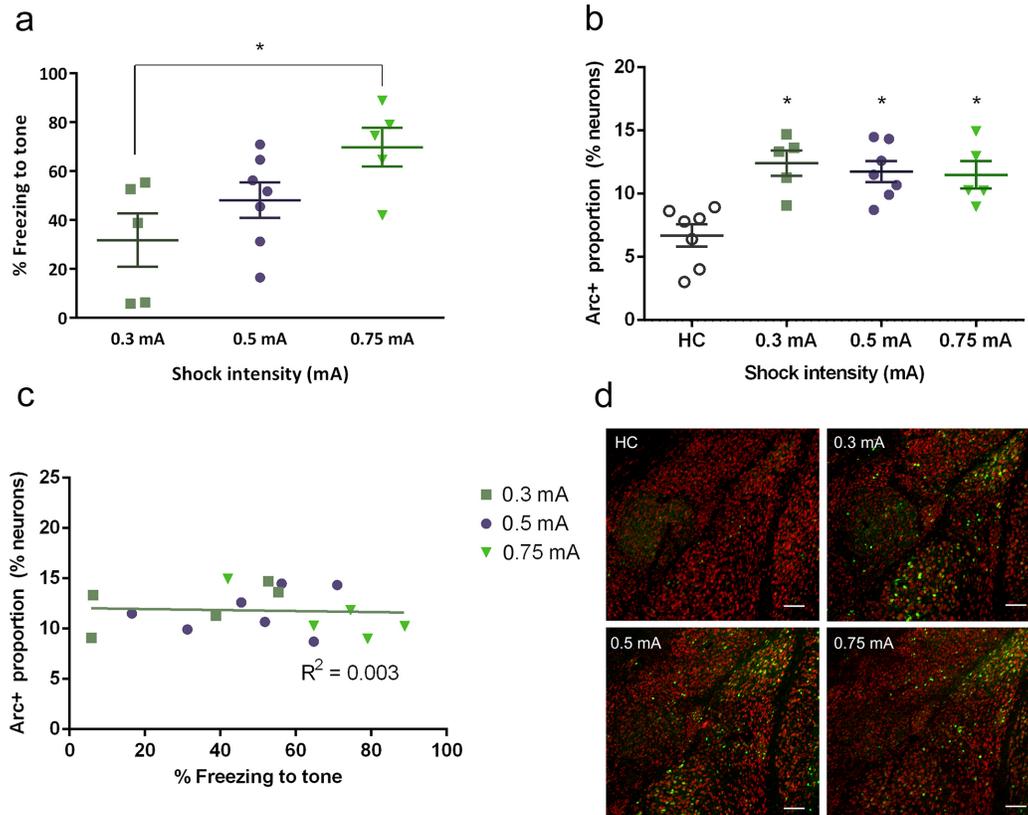
Figure 4.2: **Consistent proportion of Arc+ neurons in the LA following auditory fear retrieval despite varying degrees of freezing.** (a) Increasing shock intensity during training produced higher levels of freezing during the test, consistent with greater fear memory. (b) The proportion of Arc+ neurons was increased following testing, but no difference in Arc+ proportion was observed between groups trained with varying intensities of shock. (c) No relationship was seen between time spent freezing to tone and Arc+ proportion. (d) Examples of Arc (green) and NeuN (red) staining in the LA of HC mice and those treated with varying shock intensity. Scale bars, 50 m. (HC, n = 7; 0.3 mA footshock intensity, n = 5; n = 7, 0.5 mA; n = 5, 0.75 mA). * $P < 0.05$ with Tukey's post hoc test for multiple comparisons. Data is expressed as mean $\pm$ sem.

tone (Figure 4.3a; $F(2, 7) = 5.30$, $P < 0.05$). However, there was no difference in the proportion of Arc^+ cells between any groups exposed to fear conditioning (Figure 4.3b; $F(4, 10) = 0.99$, $P > 0.05$). Furthermore, no difference was observed in the number of Arc^+ neurons between brains examined after training or testing. All groups that received training differed from the immediate shock control, which exhibited very low Arc signal (all groups vs immediate shock, $P < 0.001$). No relationship was observed between Arc^+ proportion and percentage time spent freezing (Figure 4.3c, $R2 = 0.096$), suggesting that the proportion of LA neurons that are strongly activated during memory encoding and retrieval 24 hr later is unrelated to the strength of learned associations.

It is worth mentioning that the average Arc^+ proportion observed in trained groups in this experiment differs from that in the previous experiment (12-15%). This may be due to the fact that the techniques used to visualized the engram, immunohistochemistry and catFISH, differ significantly in the multiple ways, including how tissue is collected, how samples are sectioned, and what is being labeled. Thus, attempting to compare these data sets side by side can be misleading.

4.4 Inhibition of PV$^+$ interneurons during conditioning increases the size of the lateral amygdala engram

In both the LA and the basal amygdala (BA), local inhibition is predominantly mediated by GABAergic parvalbumin (PV) interneurons (Ehrlich et al., 2009; Spampanato et al., 2011), which form a broad, inter-connected inhibitory network that responds to sensory input (Szinyei et al., 2000) and tightly controls
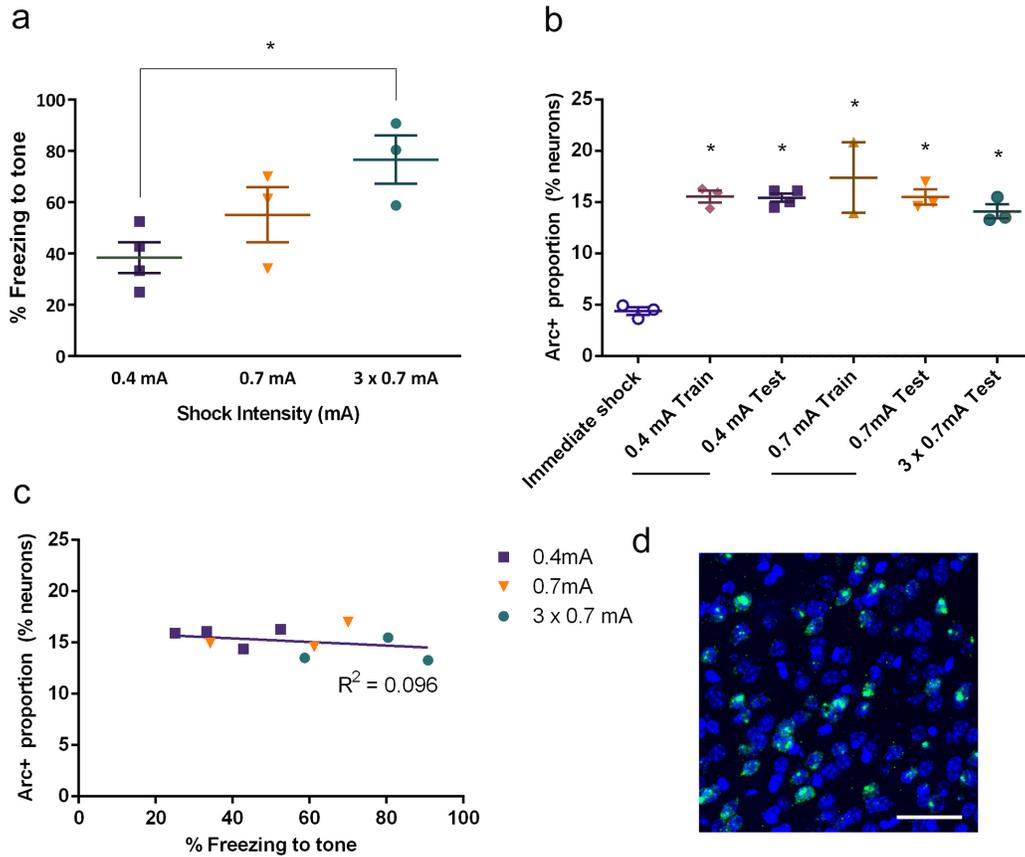
Figure 4.3: **Consistent proportion of cells expressing *Arc* mRNA in the LA following auditory fear retrieval despite varying degrees of freezing** (a) Increased shock intensity produced greater time spent freezing to tone during testing. (b) No difference in *Arc+* proportion was observed between groups trained with varying intensities of shock (c) No relationship was seen between time spent freezing to tone and *Arc+* proportion within mice. (d) Example of catFISH staining for *Arc* (green) and Hoescht (blue) in the LA. Scale bar, 50 m. (n = 3, immediate shock; n = 3, 0.4 mA Train; n = 4, 0.4 mA Test; n = 2, 0.7 mA Train; n = 3, 0.7 mA Test; n = 3, 3 x 0.7 mA Test). * P < 0.05 with Tukey's post hoc test for multiple comparisons. Data is expressed as mean ± sem.

the activity of LA principal neurons through plastic, perisomatic synapses (Freund & Katona, 2007; Trouche et al., 2013). Because of this, PV^+ interneurons represent a promising candidate for mediating inhibitory interactions between principal neurons. To determine whether PV^+ interneurons play a role in constraining the size of the LA engram, we used inhibitory DREADDS (designer receptors exclusively activated by designer drugs) to silence their activity during training. DREADDS are modified G-protein coupled receptors activated by the synthetic ligand clozapine-N-oxide (CNO) that allow neuronal activity to be modulated over prolonged periods (Nichols & Roth, 2009). The inhibitory DREADD hM4D was selectively targeted to PV-containing cells in the amygdala through the use of a Cre-dependent adeno-associated viral vector (AAV-DIO-hM4Di-mCherry) infused into the amygdala in transgenic mice (PV-Cre knockin). This technique ensured that an inhibitory receptor was expressed only in PV^+ cells of the amygdala (Tanahira et al., 2009). CNO or a vehicle (VEH) control was systemically administered 1 hr before auditory fear conditioning with a 0.5 mA footshock to inhibit PV^+ cells during memory encoding. The next day, animals were sacrificed 90 min after testing in a novel context and immunohistochemistry for Arc protein was performed.

Inhibiting PV^+ interneurons during auditory fear conditioning increased the number of active neurons in the LA following fear retrieval. No difference in freezing behavior between animals treated with CNO or VEH (Figure 4.4a; unpaired t test, $P > 0.5$) was observed, but mice treated with CNO displayed an increased proportion of Arc^+ cells in the LA (Figure 4.4b; $P < 0.01$; Mann-Whitney U test, $P < 0.05$). No relationship was observed between percentage of time spent freezing and Arc^+ proportion (Figure 4.4c; CNO, R2 = 0.015; VEH R2 < 0.001). Together, these findings suggest that reduced PV^+ interneuron

activity during auditory fear conditioning may increase the number of cells recruited to the engram without influencing memory strength.

The LA can be subdivided into distinct subnuclei on the basis of extra-amygdalar connectivity (Romanski et al., 1993; P. Sah et al., 2003). To explore whether the effect of PV inhibition varied between these subnuclei, stereological estimates of Arc$^+$ proportion in PV-Cre mice treated with CNO or VEH were divided among the dorsal (LAd), ventral-medial (LAvm), and ventral-lateral (LAvl) regions of the LA (Figure 4.5a). This analysis revealed that Arc$^+$ proportion varied between these regions in animals treated with CNO (Figure 4.5B; $P < 0.05$), but not in the VEH group ($P > 0.05$). Furthermore, the LAd, which had the greatest proportion of Arc$^+$ cells in CNO-treated animals, was also the only subnucleus in which we observed a difference following PV$^+$ inhibition (unpaired t test, $P < 0.01$). These findings suggest that the increase in the size of the engram we observed could have been driven exclusively by changes in the recruitment of LAd neurons.
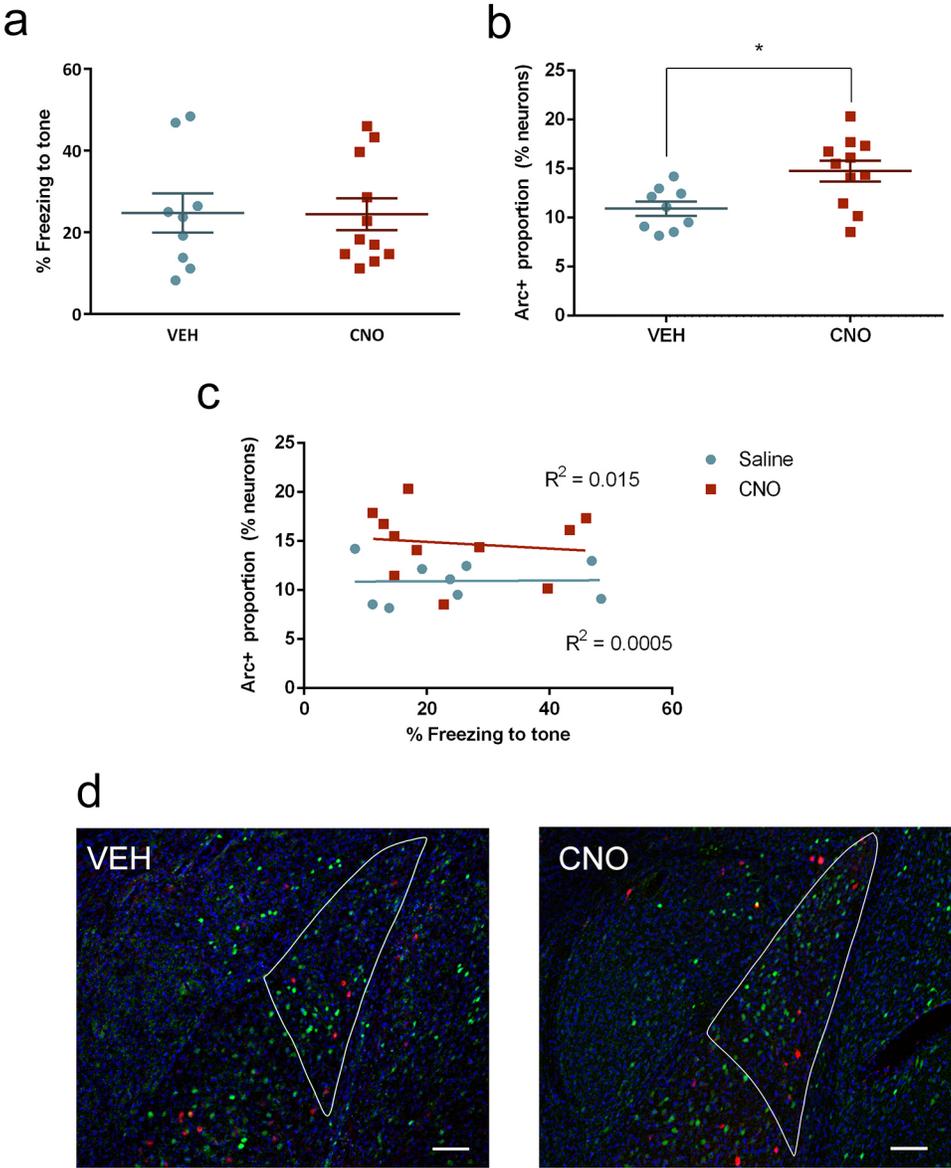
Figure 4.4: **PV+ inhibition during auditory fear conditioning increases the proportion of Arc+ neurons during subsequent retrieval.** (a) Both groups treated with CNO and VEH showed similar freezing to tone. (b) The proportion of Arc+ neurons in the LA was greater in mice treated with CNO. (c) No relationship was seen between freezing behaviour and Arc+ proportion in mice treated with VEH or CNO. (d) Representative images of sections with Arc (green), DAPI (blue), and virally infected PV$^+$ interneurons expressing mCherry (red). The LA is outlined. (n = 9, VEH; n = 11, CNO). * P < 0.05. Data is expressed as mean $\pm$ sem.
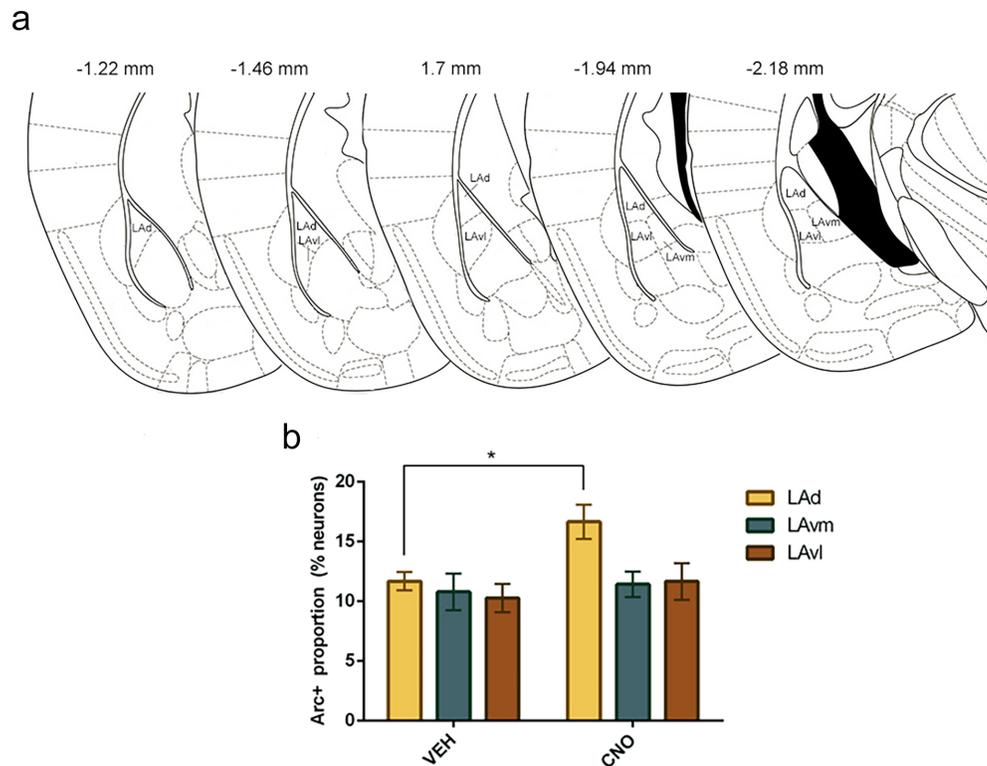
Figure 4.5: **Inhibition of PV+ cells increases the proportion of Arc+ cells in the LAd** (a) Anatomy of the amygdala from anterior (-1.22 mm from bregma) to posterior (-2.18 mm) illustrating the distinction between the LAd, LAvm, and LAvl. Adapted from Paxinos and Franklin. (b) The LAd contained a higher proportion of Arc+ cells in animals treated with CNO and was the only subnucleus in which a difference in Arc+ proportion was observed when compared to VEH controls (n = 9, VEH; n = 11, CNO). * P < 0.05 with unpaired Student's t test. Data is expressed as mean $\pm$ sem.

Chapter 5

Discussion

5.1 Summary of findings

Memories are thought to be represented in the brain by discrete biophysical changes (Eichenbaum, 2016; Josselyn et al., 2015; Tonegawa et al., 2015). These changes, collectively referred to as an engram, are widely distributed throughout the brain and include a range of alterations, from epigenetics to synaptic connectivity and neural excitability. Although little is known about how information is encoded within engrams (Eichenbaum, 2016), recent studies have shown that neurons active during learning are recruited into sparse ensembles whose reactivation at later time points is necessary and sufficient for memory recall (Han et al., 2009; J. Kim et al., 2013b; Liu et al., 2012). In the amygdala, these ensembles are primarily composed of those cells that were most strongly activated by the stimuli present during learning, with the results of some studies indicating that the size of these neuronal ensembles is constrained by mechanisms that allocate only a proportion of the most excitable cells to the engram (Han et al.,

2007; Yiu et al., 2014). Theoretical studies investigating this process of competitive memory allocation have suggested that this constraint facilitates the implementation of efficient, specific memory representations (Feng et al., 2016). However, empirical evidence for the mechanisms that mediate this constraint has not yet been collected.

In this work, I have examined the process of engram formation in the LA and attempted to prove that inhibitory interneurons constrain the number of neurons allocated to the engram to a sparse trace. First, I demonstrated that the immediate early gene Arc can be used to visualize neurons allocated to the engram of an auditory fear memory in the LA. Second, I demonstrated that the LA component of an engram remains stable despite variations in the strength of the memory. Lastly, I tested whether inhibitory PV^+ interneurons in the amygdala are involved in memory allocation by inhibiting their activity during auditory fear conditioning. I observed that the inhibition of PV^+ interneurons increased the size of the LA engram, confirming predictions made in modeling studies regarding the role of inhibitory interneurons in mediating a process of competition between neurons during memory encoding that constrains the engram to a sparse population.

5.2   The use of IEGs to identify the engram

There is a wide variety of techniques used in neuroscience to measure the activity of the brain. One such technique involves the use of immediate early genes (IEGs). IEG analysis allows researchers that study the brain in fine morphological detail with fixed tissue histology and microscopy to obtain a rough estimate

of neural activity by exploiting the fact that many genes are reliably transcribed and translated into protein following intense neural activity. When examining fixed brain tissue specimens, immunohistochemical staining for these genes permits the cells that were most active in the period of time directly before fixation to be identified.

Genes used for IEG analysis share a number of common properties. Crucially, they are all expressed constitutively at very low levels, but increase their expression rapidly and dramatically following the types of high frequency activity associated with synaptic plasticity and neural encoding (J. F. Guzowski et al., 1999). The functions of the proteins these genes encode are diverse, but many are involved in synaptic plasticity and other processes that allow neurons to modify their function in response to external input (John F. Guzowski et al., 2005). Genes used for IEGs also tend to be highly conserved across species, making them appropriate targets for study in a wide variety of animal models.

Activity-regulated cytoskeletal-associated protein (Arc) is an IEG that has the specific advantage of being selectively expressed only in excitatory principal neurons (Vazdarjanova et al., 2006). Arc interacts with the synaptic machinery of receptor endocytosis to metaplastically regulate synapses that have recently undergone long-term potentiation (Bramham et al., 2010), and although its function has not been fully described, it has been used in a large body of work to identify neurons allocated to the engram in the LA (Gouty-Colomer et al., 2015; Reijmers et al., 2007), hippocampus (Liu et al., 2012; Tayler et al., 2013).

5.3 Arc expression increases during training and testing

In order to validate the use of Arc to measure engrams in this study, immunohistochemistry for Arc was performed after either auditory fear training or testing the density of Arc$^+$ cells compared to treatment groups in which no training was given (homecage) or in non-learning control groups (tone only, context only, immediate shock). In all but the intercalated cell masses, the density of Arc$^+$ cells was greater in animals sacrificed after either training or testing than any control groups, suggesting that the increase in Arc across the LA seen during training and testing was specifically due to the activation of neuronal ensembles involved in the encoding and retrieval of an associative memory. It was also determined that the density of Arc$^+$ cells was equivalent between testing and training conditions within all brain regions, consistent with the hypothesis that similar patterns of activity are induced in the brain during both encoding and retrieval (Reijmers et al., 2007; Tayler et al., 2013).

The increase in Arc seen throughout the amygdala is highly consistent with the current understanding of the storage and retrieval of fearful memories. Although the LA, which acts as the primary sensory interface of the amygdala contains the essential components of the engram for encoding sensory associations, activity in the BA and CeA is thought to underlie the inclusion of contextual details and fear expression, respectively (Tovote et al., 2015). The lack of Arc response seen in the intercalated cell masses is likely due to the fact that they are primarily composed of GABAergic interneurons (Ehrlich et al., 2009), which do not express Arc (Vazdarjanova et al., 2006).

5.4   Stereological counting to measure the engram

In order to accurately measure the proportion of neurons allocated to an engram, unbiased, systematic stereological procedures were used. By systematically drawing samples from a large population, the practice of stereology allows geometrical quantities to be estimated without sacrificing accuracy or efficiency. When attempting to measure the number of cells in a region of the brain, exhaustive counting of dense cellular populations becomes formidably time-consuming. In addition, sectioning tissue often leads to cell fragments, which can produce estimates of cell number that differ from the true number of cells present. Stereological techniques overcome these issues by drawing samples for counting from a large total volume in a statistically unbiased way, so the density of cells within a smaller counted area can be extrapolated to represent the larger volume.

To obtain an accurate ratio of Arc^+ to total neurons in the LA, serial sections were analyzed using optical fractionation. Optical fractionation begins with the collection of serial sections throughout the region of interest, followed by the systematic random distribution of three-dimensional counting frames within those sections. The distance between sections and the size and density of counting frames is determined before the experiment begins in order to achieve an appropriate level of statistical accuracy (West et al., 1991).

5.5   Memory retrieval activates a constant proportion of cells in the amygdala despite varying memory strength

The hypothesis that competitive inhibitory interactions enforce sparsity within the LA engram suggests that all memories will be distributed among neural ensembles of the same size, regardless of the content or intensity of those memories. One way to test whether this is the case is to manipulate memory strength (ie. the distinctiveness and/or intensity of learned associations) and investigate whether this alters the number of neurons allocated to the engram. Because the strength of Pavlovian conditioning memories is a function of the valence of the US and the salience of the CS (Itzhak et al., 2014), manipulationg the intensity of US footshock delivered during auditory fear conditioning was explored as a method to increase fear memory strength. Indeed, increasing footshock was associated with greater time spent freezing to the tone 24 hrs after training.

Stereological counting of the proportion of neurons active during retrieval did not reveal a relationship between memory strength and the size of the engram. All animals that showed elevated freezing during tone presentation displayed a higher degree of Arc labeling in the LA compared to animals that remained in their homecage, indicating that successful retrieval of a fear association was associated with reactivation of a small subset of LA neurons. However, there was no difference in the size of this population between groups trained with low (0.3 mA), medium (0.5 mA), or high (0.75 mA) footshock. Furthermore, no relationship was seen between an animal's degree of freezing to the tone and the proportion of Arc^+ cells in the LA. These results suggest that all auditory fear engrams are stored in similarly sized populations of neurons in the LA (~12% in these experiments).

These findings do not preclude the existence of differences between the LA engrams of strong or weak memories. Rather, they only suggest that it is the number of neurons involved that remains constant. There are many components of the engram that the approach taken in this study was unable to detect, so it is possible that characteristics of similarly-sized memory ensembles may vary along these other dimensions. Indeed, it has been shown that engram neurons are more intensely activated during the retrieval of strong memories (Gouty-Colomer et al., 2015). Furthermore, it is possible that the size of engrams in the LA may change over time and that the number of neurons activated during retrieval may be reduced due to extinction or consolidation. One previous study that reported a positive relationship between the strength of auditory fear memory and the number of reactivated neurons in the LA did so only after extinction training (Reijmers et al., 2007). Thus, it is possible that the number of reactivated neurons in the engram may bear a significant relation to memory strength in conditions in which the memory trace has been degraded or suppressed. In these cases, a reduced number of active neurons might represent a failure to reactivate the entire neuronal ensemble associated with the engram.

5.6 messenger RNA (mRNA) as a time sensitive marker of neuronal activity

The primary limitation of protein-based IEG techniques for identifying the engram is an extremely poor temporal resolution. Protein expression occurs over a period of minutes and hours, whereas the sensory-behavioural events that IEG approaches attempt to investigate occur often over a time period of seconds.

Thus protein-based labeling often captures not only neurons active during a be-
havioural epoch of interest, but also those active during a broad period of time
before and after. This drawback has been suggested to underlie some of the
surprisingly low estimates of overlap between neural populations active during
encoding and retrieval in previous studies (Josselyn et al., 2015).

One technique that can be used to overcome the limited temporal resolu-
tion of protein-based IEG analysis is cellular compartmental analysis of tempo-
ral activity by fluorescence in situ hybridization (catFISH), which labels mRNA
of IEGs. This technique benefits from the fact that mRNA transcription in the
nucleus begins much earlier than changes in protein expression after neural activ-
ity and has been shown to be able to identify increases in Arc mRNA associated
with high frequency activity in as little as 2 mins (John F. Guzowski et al., 2005
). This stands in contrast to Arc protein expression, which does not peak un-
til 60-90 minutes after activity. The catFISH technique also allows multiple
time points of activity to be detected due to the fact that mRNA transcripts
produced in the nucleus transit to the dendrites to be expressed: dendritic Arc
indicates activity between 20-60 min ago while nuclear Arc indicates activity
with < 16 min.

5.7 Arc mRNA reveals sparse activation of LA neurons during fear memory retrieval

In order to demonstrate that the stable proportion of engram cells observed with
multiple shock conditions was not due to an over-inclusive labeling technique,
catFISH and stereological counting was performed to determine the proportion

of neurons activated within a narrow time window around encoding or retrieval of an auditory fear memory. This investigation produced a similar result as the previous investigation involving protein-based IEGs. Increased shock intensity was associated with more strongly expressed memories, but there was no difference in the size of the LA component of the engram between any group during training or testing. This suggests that, even when correcting for the temporal inaccuracy of protein-based IEGs, the proportion of neurons allocated to the engram in memories of different strengths remains constant.

Interestingly, despite the narrow temporal window afforded by catFISH, a similar, slightly greater, proportion of engram neurons (15%) was recorded in comparison to experiments involving protein-based IEG, even when only cells with intranuclear fluorescence were counted. It is unclear whether the smaller time window of capture was offset by a lower threshold of detection with cat-FISH, or whether the previous protein based IEG analysis was simply accurate enough. However, the relative similarity in estimates of the engram population between this mRNA-based approach, protein-based results, and the multitude of other studies on engram size (An et al., 2012; Ghosh & Chattarji, 2015; Gouty-Colomer et al., 2015; Herry et al., 2008; Quirk et al., 1995) provide further support for the hypothesis that memory traces are actively constrained to a sparse population and validation of this experimental approach involving Arc.

5.8   DREADDS to manipulate cellular excitability

In order to selectively modulate the activity of PV$^+$ interneurons during the acquisition of fear memories, DREADDs were targeted to PV-expressing neurons

with a Cre-dependent viral vector. DREADDs are modified G-protein coupled receptors that are activated by a specific synthetic ligand, clozapine-N-oxide (CNO), that otherwise has no activity in the body (Armbruster et al., 2007). Several varieties of DREADDs have been created that have different functions on neural activity. The DREADD subtype used in this investigation, hM4Di, is an inhibitory receptor that decreases cellular activity by lowering the activity of cAMP, reducing potassium channel conductance, and inhibiting presynaptic neurotransmitter release.

DREADDs allow neural activity to be modulated over a time-scale that is both appropriate and highly convenient for behavioural neurobiological experiments. When CNO is injected intraperitoneally, DREADD activation begins within minutes, peaks after 50 minutes, and dissipates several hours later (Alexander et al., 2009). This provides a sufficiently large time window for behavioural experiments to be performed while also ensuring that experiments performed the following day are free of the effects of lingering DREADD activity. The consequences of the prolonged neural modulation that DREADDs provide during the few hours that CNO is present are not entirely understood, however. Dramatic differences in outcomes between studies that used prolonged, pharmacological manipulations and precise optogenetic approaches have been identified (Goshen et al., 2011), and it may be possible that the brain is able to compensate for the effect of DREADDs in some way.

5.9 Inhibition of PV$^+$ interneurons increases the size of the LA engram

Computational modeling of memory formation in the LA has suggested that competitive interactions mediated by inhibitory interneurons are necessary for maintaining the size of memory traces despite variations in excitability (Feng et al., 2016). In order to empirically test this hypothesis, DREADDs were selectively targeted to PV-containing interneurons in the amygdala with the use of PV-Cre transgenic mice and a Cre-dependent viral vector. This allowed a proportion of PV$^+$ interneurons to be inhibited during auditory fear conditioning to assess their role in constraining the size of the engram in the LA. Of the many interneuron subtypes present, we focused on PV$^+$ interneurons because of their suggested role in the regulation of principal neuron excitability (Spampanato et al., 2011) and the modulation of fear memories (Trouche et al., 2013).

Remarkably, inhibiting PV$^+$ cells during training increased the number of neurons allocated to the engram (Figure 4.4), suggesting that they may play some role in the mechanisms that govern the formation of neuronal ensembles during memory encoding. This increase was modest (11 –> 15%), but suggests that the LA component of the fear engram was 35% larger when formed under conditions of reduced PV$^+$ activity. Although this investigation is unable to describe exactly how this expansion in the memory trace occurred, these results corroborate modeling studies regarding the importance of inhibitory interneurons in making memory allocation a competitive process (Feng et al., 2016; Dongbeom Kim et al., 2013b). The most recent study proposed that decreases in the plasticity of inhibitory connections would lead to an increase in the number of cells that undergo plasticity during fear conditioning. Reducing

the plasticity of inhibitory interneurons in the model during learning lead to an 35% increase in the number of excitatory neurons that underwent plasticity, a result that is strikingly similar to what I observe in my experiment.

These results do not preclude the possibility that other interneuronal sub-types may also be involved in constraining the size of the engram (Stefanelli et al., 2016). Indeed, it may be possible that this increase in the size of the engram observed during PV^+ interneuron inhibition was due to a disruption of the PV^+ and SOM^+ microcircuits that shape the flow of sensory input (Wolff et al., 2014).

The lack of untrained controls in the PV^+ inhibition experiment makes it impossible to firmly conclude that the increase in the number of activated cells during testing was specifically due to the increased recruitment of tone-responsive neurons to the engram. It remains to be seen whether short-term PV^+ interneuron inhibition alone, without training, is capable of causing long-lasting changes that may influence the number of Arc^+ cells during a test 24 hr later. However, these results do suggest that interneurons play some role in the plastic processes that occur during fear conditioning in the LA.

5.10 PV^+ inhibition-related engram increases are specifically located in the LAd

In order to understand the nature of these findings in light of the anatomical distinctions that have been proposed between the LA's dorsal (LAd), ventral-medial (LAvm), and ventral-lateral (LAvl) subnuclei (Romanski et al., 1993; P. Sah et al., 2003), the stereological results of the experiment involving PV^+ cell

inhibition were divided up between subnuclei to obtain independent estimates of Arc^+ proportion. These three subdivisions of the LA have been shown to receive different forms of innervation from the rest of the brain, with the LAd receiving the majority of fast, thalamic, auditory projections and the ventral regions receiving slower, cortical inputs (Romanski et al., 1993). Thus, it is possible that salient sensory information encoded during auditory fear conditioning may be differentially allocated across these regions.

Interestingly, the proportion of Arc^+ neurons active during memory retrieval was significantly higher in the LAd of animals in which PV^+ interneurons were inhibited during training (Figure 4.5). Furthermore, the effects of PV inhibition in the LAd accounted for the nearly all of the increased size of the engram. It is possible that PV^+ interneurons play a greater role in regulating the excitability of neurons in the LAd rather than the LAvm or LAvl. There was no difference in the density of engram cells between subdivisions in control animals. Thus, it is possible that, under normal conditions, neurons throughout the LA are allocated to the engram of an auditory fear memory with equal probability. It is unclear whether the increase in the proportion of engram cells in the LAd is due to greater innervation by auditory thalamic projections (Romanski et al., 1993), different plastic properties (Repa et al., 2001), or tighter regulation from inhibitory interneurons (Dongbeom Kim et al., 2013d). However, these results do provide strong evidence for significant functional differences between subnuclei of the amygdala with regard to the composition of engrams that deserves further exploration.

Chapter 6

Conclusion

6.1 How could interneurons shape the size of the engram?

Although the results of this thesis are unable to completely describe how PV^+ interneurons modulate the size of the LA memory trace, they support a model in which the size of the engram is shaped by competitive inhibitory interactions between principal neurons (Dongbeom Kim et al., 2013b). In this model, as sensory input flows into the LA during fear conditioning, principal neurons that receive sensory innervation are excited, with some neurons being more excited than others based on variations in intrinsic excitability (likely related to differences in CREB expression that fluctuate over time). The most highly excitable neurons in the network that share excitatory connections will further increase each other's activity, leading to the autoassociative induction of Hebbian plasticity. These 'winners' (neurons that are excited above plasticity thresholds during training), will be selectively allocated to the engram. However, 'winners' will also activate a population of inhibitory interneurons (many of them PV^+)

that dampen the activity of many neurons in the network, converting many would-be 'winners' to 'losers', thereby reducing the size of the engram. In this model, the pre-existing set of connections between LA neurons is important, since those that mostly receive direct, excitatory connections from the most excitable 'winners' will be much more likely to be allocated to the engram than those that receive primarily inhibitory connections from interneurons activated by those 'winners.'

In addition to constraining the size of the LA engram, this model also describes how the size of the engram remains constant despite variations in training intensity. During weak training conditions, the low intensity of US inputs brings only a few neurons above the threshold of activity necessary for plasticity. Thus, the engram population during weak training conditions is dependent primarily upon variations in excitability between principal neurons (Figure 6.1a). However, as the intensity of training conditions increase, stronger sensory input leads to a greater probability that more cells will be excited above plasticity thresholds. In these cases, a corresponding elevation in the activity of the inhibitory interneuronal population compensates for this increased excitation and restricts the number of 'winners' (Figure 6.1b).

Any reduction in the activity of PV^+ interneurons during fear conditioning, such as was performed in the DREADD work, would be expected to increase the size of the engram following this model. The inhibition of PV^+ interneurons impairs their ability to dampen the excitability of the many principal neurons that respond to sensory input. Without a functioning system to restrict the number of 'winners', fear memories formed in the LA will involve a greater number of neurons, as those that would ordinarily be made 'losers' by competitive inhibitory interactions are now free to cross plasticity thresholds and join

the engram.
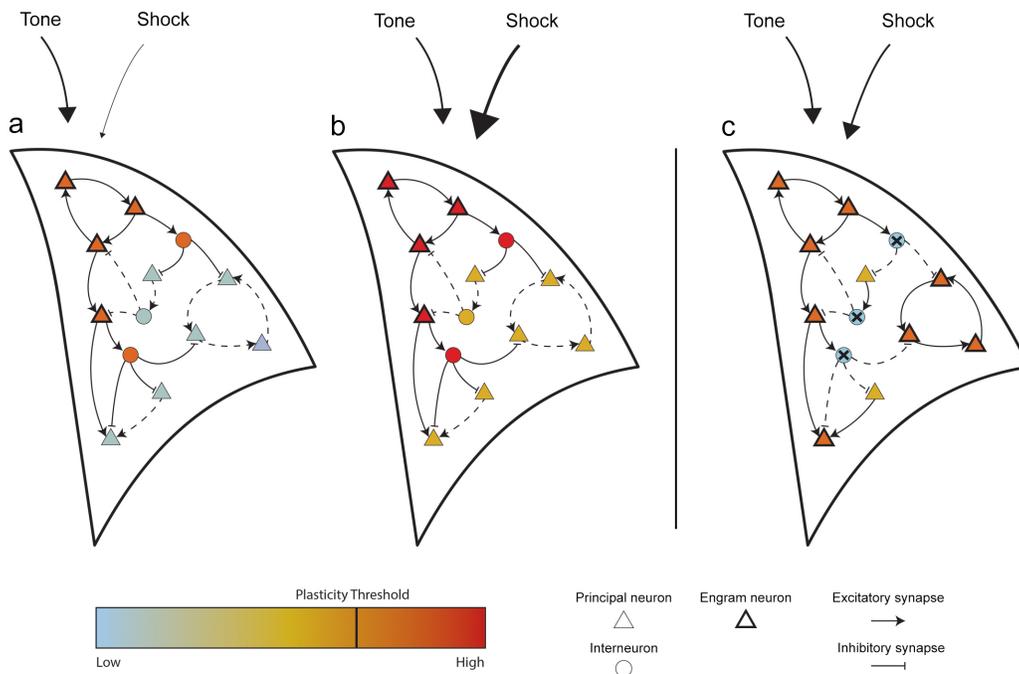
Figure 6.1: **Model of inhibitory control of the engram** Diagram of the LA during auditory fear memory formation in three different conditions. (a) Weak training with low intensity shock leads to only the most excitable and reciprocally connected neurons becoming activated above the plasticity threshold and selected to the engram. (b) Strong training leads to a greater activity in all the neurons in the LA. However, the activity of principal neurons that receive disynaptic, inhbitory projections from the most excitable subset are inhibited the plasticity threshold. (c) When PV[+] interneurons are inhibited, more principal neurons are highly active, leading to an increase in the size of the engram. All principal neurons are assumed to receive both tone and shock inputs.

6.2   Future directions

Although this study provides interesting early evidence for the involvement of interneurons in memory allocation, there are many more unanswered questions remaining. Principal among them is the question of whether only PV[+] interneu-

rons influence the size of engram populations or whether other interneuronal subtypes are involved as well. Indeed, it has recently been suggested that, in the dentate gyrus of the hippocampus, SOM^+ interneurons in particular play a central role in shaping the size of memory ensembles (Stefanelli et al., 2016). There are many functional and anatomical differences between the dentate gyrus and the LA, including the relative prevalence of interneuronal subtypes (Amaral et al., 2007; Houser, 2007). However, it is possible that SOM^+ interneurons may play a role in the LA as well. Investigating whether inhibition of SOM^+ interneurons during fear conditioning might lead to similar expansions in the engram proportion will be the first step in expanding our understanding of the role of distinct interneuronal subtypes in memory allocation.

Future understanding of the ideas behind this investigation would also benefit from a more detailed characterization of PV^+ interneuron activity during memory encoding and retrieval. For instance, experiments in which PV^+ cells were excited rather than inhibited, perhaps through the use of excitatory DREADDs, could be performed in order to determine whether memory traces could be constrained even further. Furthermore, electrophysiological recording studies could be performed to determine whether interneurons fire more strongly during conditions of strong memory and match the patterns of activity proposed under the hypothesis of competitive inhibitory interactions. It is also possible that analyzing the characteristics of inhibitory potentials received from principal cells during auditory fear condition could allow the source of inhibition to be determined (ie. from PV^+ or SOM^+ cells).

Theoretical studies on memory allocation in the LA propose that constraining the memory trace is advantageous by promoting a balance between the specificity and robustness of learned associations (D. Kim et al., 2015). Al-

though our study did not reveal any relationship between memory strength and the size of the engram, it would be interesting to determine whether memories that are stored in more neurons differ in less obvious ways. For example, exposing fear conditioned mice to tones different from those they were exposed to during training could allow an effect of engram size on memory generalization to be detected. Additionally, the resistance of memories stored in more neurons to natural forgetting over time or extinction could be explored. Both of these findings may have important implications in the study of anxiety disorders and PTSD, which are often conceived of as symptoms of over-generalized and hyper-persistent threat detection (Lapiz-Bluhm & Peterson, 2014; Tovote et al., 2015).

Although there was a cursory investigation into the spatial location of engram cells within the LA, the significant functional and anatomical differences that have been identified between subnuclei of the LA merits further investigation. The finding that the LAd in particular showed the greatest increase in engram cells is interesting considering several lines of evidence that implicate the most dorsal aspects of the LA in receiving the greatest amount of sensory innervation and mediating the earliest, most transient forms of plasticity during fear conditioning (Repa et al., 2001, Romanski et al. (1993)).

It could also be fruitful to use techniques such as CLARITY, which involves removing lipids from brain tissue to produce samples that are capable of being imaged in three dimensions (Tomer et al., 2014), to analyze the distribution of cells throughout the entire volume of the amygdala without damaging the tissue or introducing variability that prevents accurately defining the subnuclei. Such an investigation could also be combined with staining for molecular markers or labeling of sensory projections that would allow the boundaries be-

tween subnuclei to be defined and the relationship between the input a neuron receives and its likelihood to be allocated to the engram to be examined. With such an experimental approach, it would also be possible to examine the spatial distribution of engram neurons throughout the amygdala to uncover patterns that might be important for understanding how information is encoded in these neuronal ensembles. Indeed, if the LA does implement sparse distributed memory, memories that share similar features (eg. sensory modality, context, time of occurrence) should be stored in overlapping groups of neurons. It may one day be possible to identify how these overlapping subsets of engram contribute to shared content and reverse engineer meaning from engrams.

6.3   Conclusion

Inspired by previous discoveries from many labs and the conclusions made by computational researchers to explain those results, I have endeavored in this thesis to test the hypothesis that inhibitory neurons shape how memories are encoded in the brain. My primary finding is that silencing PV^+ inhibitory interneurons in the lateral amygdala during learning increases the number of neurons that are active during retrieval of a conditioned fear memory. Although this investigation is far from the wide-ranging body of work that will be necessary to characterize the role interneurons play in shaping engrams, I believe this advances our understanding of the conserved mechanisms of memory encoding.

Validating predictions made by theorists not only provides further insights into how the brain works, it also confirms the accuracy and usefulness of theoretical models themselves. As the field of neuroscience progresses, these compu-

tational models will only become more and more useful. After all, the brain is fundamentally a computational organ. It is my opinion that we are still awaiting the great breakthroughs in neuroscience that will allow us to enhance human performance and cure neurological illness. However, to make these breakthroughs we will need honest, and efficient collaboration between experimentalists and theoreticians. This may involve diverse skill sets within labs, more interdisciplinary training, and the open sharing of data and code, all of which will allow both sides to be more productive. Indeed, there will soon be so much expertise required in both the technical and theoretical aspects of neuroscience that that it will be the only way anything will ever get done.

References

Ahmad, S., & Hawkins, J. (2015). Properties of Sparse Distributed Representations and their Application to Hierarchical Temporal Memory. *ArXiv:1503.07469 [Cs, Q-Bio]*.

Alexander, G. M., Rogan, S. C., Abbas, A. I., Armbruster, B. N., Pei, Y., Allen, J. A., … Roth, B. L. (2009). Remote control of neuronal activity in transgenic mice expressing evolved G protein-coupled receptors. *Neuron*, *63*(1), 27–39.

Amano, T., Duvarci, S., Popa, D., & Paré, D. (2011). The Fear Circuit Revisited: Contributions of the Basal Amygdala Nuclei to Conditioned Fear. *J. Neurosci.*, *31*(43), 15481–15489.

Amaral, D. G., Scharfman, H. E., & Lavenex, P. (2007). The dentate gyrus: Fundamental neuroanatomical organization (dentate gyrus for dummies). *Prog Brain Res*, *163*, 3–22.

An, B., Hong, I., & Choi, S. (2012). Long-Term Neural Correlates of Reversible Fear Learning in the Lateral Amygdala. *J. Neurosci.*, *32*(47), 16845–16856.

Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S., & Roth, B. L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl. Acad. Sci. U.S.A.*, *104*(12), 5163–5168.

Bach, D. R., Weiskopf, N., & Dolan, R. J. (2011). A Stable Sparse Fear Memory Trace in Human Amygdala. *J. Neurosci.*, *31*(25), 9383–9389.

Bergstrom, H. C., & Johnson, L. R. (2014). An organization of visual and auditory fear

conditioning in the lateral amygdala. *Neurobiology of Learning and Memory*, *116*, 1–13.

Beyeler, A., Namburi, P., Glober, G. F., Simonnet, C., Calhoon, G. G., Conyers, G. F., … Tye, K. M. (2016). Divergent Routing of Positive and Negative Information from the Amygdala during Memory Retrieval. *Neuron*, *90*(2), 348–361.

Bliss, T. V., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol. (Lond.)*, *232*(2), 331–356.

Botta, P., Demmou, L., Kasugai, Y., Markovic, M., Xu, C., Fadok, J. P., … Lüthi, A. (2015). Regulating anxiety with extrasynaptic inhibition. *Nat. Neurosci.*, *18*(10), 1493–1500.

Bramham, C. R., Alme, M. N., Bittins, M., Kuipers, S. D., Nair, R. R., Pai, B., … Wibrand, K. (2010). The Arc of synaptic memory. *Exp Brain Res*, *200*(2), 125–140.

Bruce, D. (2001). Fifty years since Lashley's In search of the Engram: Refutations and conjectures. *J Hist Neurosci*, *10*(3), 308–318.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*(4466), 207–210.

Cowansage, K. K., Shuman, T., Dillingham, B. C., Chang, A., Golshani, P., & Mayford, M. (2014). Direct reactivation of a coherent neocortical memory of context. *Neuron*, *84*(2), 432–441.

Debarnot, U., Sperduti, M., Di Rienzo, F., & Guillot, A. (2014). Experts bodies, experts minds: How physical and mental training shape the brain. *Front. Hum. Neurosci.*, *8*, 280.

Denny, C. A., Kheirbek, M. A., Alba, E. L., Tanaka, K. F., Brachman, R. A., Laughman, K. B., … Hen, R. (2014). Hippocampal Memory Traces Are Differentially Modulated by Experience, Time, and Adult Neurogenesis. *Neuron*, *83*(1), 189–201.

Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subserve source memory. *Neuron*, *35*(5),

989–996.

Druckmann, S., & Chklovskii, D. B. (2012). Neuronal Circuits Underlying Persistent Representations Despite Time Varying Activity. *Current Biology*, *22*(22), 2095–2103.

Duan, W., Zhang, Y.-P., Hou, Z., Huang, C., Zhu, H., Zhang, C.-Q., & Yin, Q. (2016). Novel Insights into NeuN: From Neuronal Marker to Splicing Regulator. *Mol. Neurobiol.*, *53*(3), 1637–1647.

Ehrlich, I., Humeau, Y., Grenier, F., Ciocchi, S., Herry, C., & Lüthi, A. (2009). Amygdala inhibitory circuits and the control of fear memory. *Neuron*, *62*(6), 757–771.

Eichenbaum, H. (2016). Still searching for the engram. *Learn Behav*.

Fanselow, M. S. (1998). Pavlovian conditioning, negative feedback, and blocking: Mechanisms that regulate association formation. *Neuron*, *20*(4), 625–627.

Fanselow, M. S., & Bolles, R. C. (1979). Naloxone and shock-elicited freezing in the rat. *J Comp Physiol Psychol*, *93*(4), 736–744.

Feng, F., Samarth, P., Paré, D., & Nair, S. S. (2016). Mechanisms underlying the formation of the amygdalar fear memory trace: A computational perspective. *Neuroscience*, *322*, 370–376.

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nat Rev Neurosci*, *6*(2), 119–130.

Frankland, P. W., Josselyn, S. A., Anagnostaras, S. G., Kogan, J. H., Takahashi, E., & Silva, A. J. (2004). Consolidation of CS and US representations in associative fear conditioning. *Hippocampus*, *14*(5), 557–569.

Freund, T. F., & Katona, I. (2007). Perisomatic Inhibition. *Neuron*, *56*(1), 33–42.

Garner, A. R., Rowland, D. C., Hwang, S. Y., Baumgaertel, K., Roth, B. L., Kentros, C., & Mayford, M. (2012). Generation of a synthetic memory trace. *Science*, *335*(6075), 1513–1516.

Gdalyahu, A., Tring, E., Polack, P.-O., Gruver, R., Golshani, P., Fanselow, M. S., … Trachtenberg, J. T. (2012). Associative fear learning enhances sparse network coding in primary

sensory cortex. *Neuron*, *75*(1), 121–132.

Ghosh, S., & Chattarji, S. (2015). Neuronal encoding of the switch from specific to generalized fear. *Nature Neuroscience*, *18*(1), 112–120.

Gluck, M., & Myers, C. (2001). *Gateway to Memory: An Introduction to Neural Network Modeling of the Hippocampus and Learning.* Cambridge, Massachusetts: The MIT Press.

Goshen, I., Brodsky, M., Prakash, R., Wallace, J., Gradinaru, V., Ramakrishnan, C., & Deisseroth, K. (2011). Dynamics of retrieval strategies for remote memories. *Cell*, *147*(3), 678–689.

Gouty-Colomer, L. A., Hosseini, B., Marcelo, I. M., Schreiber, J., Slump, D. E., Yamaguchi, S., … Kushner, S. A. (2015). Arc expression identifies the lateral amygdala fear memory trace. *Mol. Psychiatry.*

Graybiel, A. M. (2005). The basal ganglia: Learning new tricks and loving it. *Current Opinion in Neurobiology*, *15*(6), 638–644.

Guzowski, J. F., McNaughton, B. L., Barnes, C. A., & Worley, P. F. (1999). Environment-specific expression of the immediate-early gene Arc in hippocampal neuronal ensembles. *Nat. Neurosci.*, *2*(12), 1120–1124.

Guzowski, J. F., Timlin, J. A., Roysam, B., McNaughton, B. L., Worley, P. F., & Barnes, C. A. (2005). Mapping behaviorally relevant neural circuits with immediate-early gene expression. *Curr. Opin. Neurobiol.*, *15*(5), 599–606.

Han, J.-H., Kushner, S. A., Yiu, A. P., Cole, C. J., Matynia, A., Brown, R. A., … Josselyn, S. A. (2007). Neuronal Competition and Selection During Memory Formation. *Science*, *316*(5823), 457–460.

Han, J.-H., Kushner, S. A., Yiu, A. P., Hsiang, H.-L. L., Buch, T., Waisman, A., … Josselyn, S. A. (2009). Selective erasure of a fear memory. *Science*, *323*(5920), 1492–1496.

Hebb, D. (1949). *The Organization of Behavior: A Neuropsychological Theory.* Wiley.

Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., & Lüthi, A. (2008). Switching on

and off fear by distinct neuronal circuits. *Nature*, *454*(7204), 600–606.

Hippenmeyer, S., Vrieseling, E., Sigrist, M., Portmann, T., Laengle, C., Ladle, D. R., & Arber, S. (2005). A Developmental Switch in the Response of DRG Neurons to ETS Transcription Factor Signaling. *PLOS Biol*, *3*(5), e159.

Houser, C. R. (2007). Interneurons of the dentate gyrus: An overview of cell types, terminal fields and neurochemical identity. *Prog. Brain Res.*, *163*, 217–232.

Hromádka, T., DeWeese, M. R., & Zador, A. M. (2008). Sparse Representation of Sounds in the Unanesthetized Auditory Cortex. *PLOS Biol*, *6*(1), e16.

Hsiang, H.-L. L., Epp, J. R., van den Oever, M. C., Yan, C., Rashid, A. J., Insel, N., … Josselyn, S. A. (2014). Manipulating a "cocaine engram" in mice. *J. Neurosci.*, *34*(42), 14115–14127.

Itzhak, Y., Perez-Lanza, D., & Liddie, S. (2014). The strength of aversive and appetitive associations and maladaptive behaviors. *IUBMB Life*, *66*(8), 559–571.

Janak, P. H., & Tye, K. M. (2015). From circuits to behaviour in the amygdala. *Nature*, *517*(7534), 284–292.

Jasnow, A. M., Ressler, K. J., Hammack, S. E., Chhatwal, J. P., & Rainnie, D. G. (2009). Distinct Subtypes of Cholecystokinin (CCK)-Containing Interneurons of the Basolateral Amygdala Identified Using a CCK Promoter-Specific Lentivirus. *Journal of Neurophysiology*, *101*(3), 1494–1506.

Jog, M. S., Kubota, Y., Connolly, C. I., Hillegaart, V., & Graybiel, A. M. (1999). Building neural representations of habits. *Science*, *286*(5445), 1745–1749.

Josselyn, S. A., Köhler, S., & Frankland, P. W. (2015). Finding the engram. *Nature Reviews Neuroscience*, *16*(9), 521–534.

Kandel, E. R. (1978). *A cell-biological approach to learning.* Society for Neuroscience.

Kanerva, P. (1988). *Sparse Distributed Memory.* Cambridge, Massachusetts: The MIT Press.

Kim, D., Paré, D., & Nair, S. S. (2013a). Assignment of Model Amygdala Neurons to the

Fear Memory Trace Depends on Competitive Synaptic Interactions. *J. Neurosci.*, *33*(36), 14354–14358.

Kim, D., Paré, D., & Nair, S. S. (2013b). Assignment of model amygdala neurons to the fear memory trace depends on competitive synaptic interactions. *J. Neurosci.*, *33*(36), 14354–14358.

Kim, D., Paré, D., & Nair, S. S. (2013c). Mechanisms contributing to the induction and storage of Pavlovian fear memories in the lateral amygdala. *Learn. Mem.*, *20*(8), 421–430.

Kim, D., Paré, D., & Nair, S. S. (2013d). Mechanisms contributing to the induction and storage of Pavlovian fear memories in the lateral amygdala. *Learn. Mem.*, *20*(8), 421–430.

Kim, D., Samarth, P., Feng, F., Pare, D., & Nair, S. S. (2015). Synaptic competition in the lateral amygdala and the stimulus specificity of conditioned fear: A biophysical modeling study. *Brain Struct Funct.*

Kim, J., Kwon, J.-T., Kim, H.-S., & Han, J.-H. (2013a). CREB and neuronal selection for memory trace. *Front Neural Circuits*, *7*, 44.

Kim, J., Kwon, J.-T., Kim, H.-S., Josselyn, S. A., & Han, J.-H. (2013b). Memory recall and modifications by activating neurons with elevated CREB. *Nature Neuroscience*, *17*(1), 65–72.

Kim, S.-Y., Adhikari, A., Lee, S. Y., Marshel, J. H., Kim, C. K., Mallory, C. S., … Deisseroth, K. (2013). Diverging neural pathways assemble a behavioural state from separable features in anxiety. *Nature*, *496*(7444), 219–223.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A Neostriatal Habit Learning System in Humans. *Science*, *273*(5280), 1399–1402.

Krashes, M. J., Koda, S., Ye, C., Rogan, S. C., Adams, A. C., Cusher, D. S., … Lowell, B. B. (2011). Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J. Clin. Invest.*, *121*(4), 1424–1428.

Krieg, D., & Triesch, J. (2014). A unifying theory of synaptic long-term plasticity based on

a sparse distribution of synaptic strength. *Front. Synaptic Neurosci, 6*, 3.

Krupa, D. J., Thompson, J. K., & Thompson, R. F. (1993). Localization of a memory trace in the mammalian brain. *Science, 260*(5110), 989–991.

Lapiz-Bluhm, M. D., & Peterson, A. L. (2014). Neurobehavioral Mechanisms of Traumatic Stress in Post-traumatic Stress Disorder. *Curr Top Behav Neurosci, 18*, 161–190.

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience, 23*(1), 155–184.

LeDoux, J. E., Cicchetti, P., Xagoraris, A., & Romanski, L. M. (1990). The lateral amygdaloid nucleus: Sensory interface of the amygdala in fear conditioning. *J. Neurosci., 10*(4), 1062–1069.

Lee, S., Kim, S.-J., Kwon, O.-B., Lee, J. H., & Kim, J.-H. (2013). Inhibitory networks of the amygdala for emotional memory. *Front Neural Circuits, 7*, 129.

Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, A., Deisseroth, K., & Tonegawa, S. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature, 484*(7394), 381–385.

Lovett-Barron, M., Turi, G. F., Kaifosh, P., Lee, P. H., Bolze, F., Sun, X.-H., … Losonczy, A. (2012). Regulation of neuronal input transformations by tunable dendritic inhibition. *Nat Neurosci, 15*(3), 423–430.

Malenka, R. C., & Bear, M. F. (2004). LTP and LTD: An embarrassment of riches. *Neuron, 44*(1), 5–21.

Markram, H., Lübke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs. *Science, 275*(5297), 213–215.

Mishkin, M. (1978). Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. *Nature, 273*(5660), 297–298.

Morris, R. G., Garrud, P., Rawlins, J. N., & O'Keefe, J. (1982). Place navigation impaired

in rats with hippocampal lesions. *Nature*, *297*(5868), 681–683.

Nichols, C. D., & Roth, B. (2009). Engineered G-protein coupled receptors are powerful tools to investigate biological processes and behaviors. *Front. Mol. Neurosci.*, *2*, 16.

Ohkawa, N., Saitoh, Y., Suzuki, A., Tsujimura, S., Murayama, E., Kosugi, S., … Inokuchi, K. (2015). Artificial Association of Pre-stored Information to Generate a Qualitatively New Memory. *Cell Reports*, *11*(2), 261–269.

Pape, H.-C. (2005). GABAergic Neurons: Gate Masters of the Amygdala, Mastered by Dopamine. *Neuron*, *48*(6), 877–879.

Paxinos, G., & Franklin, K. (2001). *The mouse brain in stereotaxic coordinates, 4th edition.* Academic Press.

Pitkanen, A. (2000). Connectivity of the rat amygdaloid complex. In *The Amygdala* (2nd ed.). Oxford University Press.

Pitkänen, A., Savander, M., Nurminen, N., & Ylinen, A. (2003). Intrinsic synaptic circuitry of the amygdala. *Ann. N. Y. Acad. Sci.*, *985*, 34–49.

Pitkänen, A., Savander, V., & LeDoux, J. E. (1997). Organization of intra-amygdaloid circuitries in the rat: An emerging framework for understanding functions of the amygdala. *Trends Neurosci.*, *20*(11), 517–523.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38.

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

Rainnie, D. G., Mania, I., Mascagni, F., & McDonald, A. J. (2006). Physiological and morphological characterization of parvalbumin-containing interneurons of the rat basolateral amygdala. *The Journal of Comparative Neurology*, *498*(1), 142–161.

Ramirez, S., Liu, X., MacDonald, C. J., Moffa, A., Zhou, J., Redondo, R. L., & Tonegawa,

S. (2015). Activating positive memory engrams suppresses depression-like behaviour. *Nature*, *522*(7556), 335–339.

Redondo, R. L., Kim, J., Arons, A. L., Ramirez, S., Liu, X., & Tonegawa, S. (2014). Bidirectional switch of the valence associated with a hippocampal contextual memory engram. *Nature*.

Reijmers, L. G., Perkins, B. L., Matsuo, N., & Mayford, M. (2007). Localization of a Stable Neural Correlate of Associative Memory. *Science*, *317*(5842), 1230–1233.

Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., & LeDoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nature Neuroscience*, *4*(7), 724–731.

Romanski, L. M., Clugnet, M. C., Bordi, F., & LeDoux, J. E. (1993). Somatosensory and auditory convergence in the lateral nucleus of the amygdala. *Behav. Neurosci.*, *107*(3), 444–450.

Rumpel, S., LeDoux, J., Zador, A., & Malinow, R. (2005). Postsynaptic Receptor Trafficking Underlying a Form of Associative Learning. *Science*, *308*(5718), 83–88.

Sacks, O. (1998). *The Man Who Mistook His Wife For A Hat: And Other Clinical Tales*. Simon and Schuster.

Sah, P., Faber, E. S. L., Armentia, M. L. D., & Power, J. (2003). The Amygdaloid Complex: Anatomy and Physiology. *Physiological Reviews*, *83*(3), 803–834.

Sanes, J. N., & Donoghue, J. P. (2000). Plasticity and primary motor cortex. *Annu. Rev. Neurosci.*, *23*, 393–415.

Santoro, A. (2015). *A computational model of systems consolidation: Optimizing memory-guided behavior using abstracted hippocampal and cortical networks* (Doctoral thesis). University of Toronto.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years

of image analysis. *Nat Meth*, *9*(7), 671–675.

Semon, R. (1921). *The Mneme*. G. Allen & Unwin.

Shi, C. J., & Cassell, M. D. (1999). Perirhinal cortex projections to the amygdaloid complex and hippocampal formation in the rat. *J. Comp. Neurol.*, *406*(3), 299–328.

Shi, C., & Davis, M. (2001). Visual pathways involved in fear conditioning measured with fear-potentiated startle: Behavioral and anatomic studies. *J. Neurosci.*, *21*(24), 9844–9855.

Silva, A. J., Kogan, J. H., Frankland, P. W., & Kida, S. (1998). CREB and memory. *Annu. Rev. Neurosci.*, *21*, 127–148.

Sorvari, H., Miettinen, R., Soininen, H., & Pitkänen, A. (1996). Parvalbumin-immunoreactive neurons make inhibitory synapses on pyramidal cells in the human amygdala: A light and electron microscopic study. *Neurosci. Lett.*, *217*(2-3), 93–96.

Sosulina, L., Graebenitz, S., & Pape, H.-C. (2010). GABAergic Interneurons in the Mouse Lateral Amygdala: A Classification Study. *Journal of Neurophysiology*, *104*(2), 617–626.

Sousa, D. (2016). *How the Brain Learns* (5th ed.). Sage Publishing.

Spampanato, J., Polepalli, J., & Sah, P. (2011). Interneurons in the basolateral amygdala. *Neuropharmacology*, *60*(5), 765–773.

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci.*, *27*, 279–306.

Stefanelli, T., Bertollini, C., Lüscher, C., Muller, D., & Mendez, P. (2016). Hippocampal Somatostatin Interneurons Control the Size of Neuronal Memory Ensembles. *Neuron*, *89*(5), 1074–1085.

Szinyei, C., Heinbockel, T., Montagne, J., & Pape, H.-C. (2000). Putative Cortical and Thalamic Inputs Elicit Convergent Excitation in a Population of GABAergic Interneurons of the Lateral Amygdala. *J. Neurosci.*, *20*(23), 8909–8915.

Tanahira, C., Higo, S., Watanabe, K., Tomioka, R., Ebihara, S., Kaneko, T., & Tamamaki, N. (2009). Parvalbumin neurons in the forebrain as revealed by parvalbumin-Cre transgenic

mice. *Neuroscience Research*, *63*(3), 213–223.

Tanaka, K. Z., Pevzner, A., Hamidi, A. B., Nakazawa, Y., Graham, J., & Wiltgen, B. J. (2014). Cortical representations are reinstated by the hippocampus during memory retrieval. *Neuron*, *84*(2), 347–354.

Tayler, K. K., Tanaka, K. Z., Reijmers, L. G., & Wiltgen, B. J. (2013). Reactivation of Neural Ensembles during the Retrieval of Recent and Remote Memory. *Current Biology*, *23*(2), 99–106.

Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: Updating the index. *Hippocampus*, *17*(12), 1158–1169.

Tomer, R., Ye, L., Hsueh, B., & Deisseroth, K. (2014). Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nat. Protocols*, *9*(7), 1682–1697.

Tonegawa, S., Liu, X., Ramirez, S., & Redondo, R. (2015). Memory Engram Cells Have Come of Age. *Neuron*, *87*(5), 918–931.

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nat. Rev. Neurosci.*, *16*(6), 317–331.

Trouche, S., Sasaki, J. M., Tu, T., & Reijmers, L. G. (2013). Fear extinction causes target-specific remodeling of perisomatic inhibitory synapses. *Neuron*, *80*(4), 1054–1065.

Truitt, W. A., Johnson, P. L., Dietrich, A. D., Fitz, S. D., & Shekhar, A. (2009). Anxiety-like behavior is modulated by a discrete subpopulation of interneurons in the basolateral amygdala. *Neuroscience*, *160*(2), 284–294.

Vazdarjanova, A., Ramirez-Amaya, V., Insel, N., Plummer, T. K., Rosi, S., Chowdhury, S., … Barnes, C. A. (2006). Spatial exploration induces ARC, a plasticity-related immediate-early gene, only in calcium/calmodulin-dependent protein kinase II-positive principal excitatory and inhibitory neurons of the rat forebrain. *J. Comp. Neurol.*, *498*(3), 317–329.

Waclaw, R. R., Ehrman, L. A., Pierani, A., & Campbell, K. (2010). Developmental Origin of the Neuronal Subtypes That Comprise the Amygdalar Fear Circuit in the Mouse. *J. Neurosci.*,

30(20), 6944–6953.

Wang, X. J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.*, *24*(8), 455–463.

Weiskrantz, L. (1956). Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *J Comp Physiol Psychol*, *49*(4), 381–391.

Weliky, M., Fiser, J., Hunt, R. H., & Wagner, D. N. (2003). Coding of natural scenes in primary visual cortex. *Neuron*, *37*(4), 703–718.

West, M. J., Slomianka, L., & Gundersen, H. J. (1991). Unbiased stereological estimation of the total number of neurons in thesubdivisions of the rat hippocampus using the optical fractionator. *Anat. Rec.*, *231*(4), 482–497.

Wixted, J. T., Squire, L. R., Jang, Y., Papesh, M. H., Goldinger, S. D., Kuhn, J. R., … Steinmetz, P. N. (2014). Sparse and distributed coding of episodic memory in neurons of the human hippocampus. *PNAS*, *111*(26), 9621–9626.

Wolff, S. B. E., Gründemann, J., Tovote, P., Krabbe, S., Jacobson, G. A., Müller, C., … Lüthi, A. (2014). Amygdala interneuron subtypes control fear learning through disinhibition. *Nature*, *509*(7501), 453–458.

Woodson, W., Farb, C. R., & Ledoux, J. E. (2000). Afferents from the auditory thalamus synapse on inhibitory interneurons in the lateral nucleus of the amygdala. *Synapse*, *38*(2), 124–137.

Yiu, A. P., Mercaldo, V., Yan, C., Richards, B., Rashid, A. J., Hsiang, H.-L. L., … Josselyn, S. A. (2014). Neurons Are Recruited to a Memory Trace Based on Relative Neuronal Excitability Immediately before Training. *Neuron*, *83*(3), 722–735.

Zelikowsky, M., Hersman, S., Chawla, M. K., Barnes, C. A., & Fanselow, M. S. (2014). Neuronal ensembles in amygdala, hippocampus, and prefrontal cortex track differential components of contextual fear. *J. Neurosci.*, *34*(25), 8462–8466.

Zhou, Y., Won, J., Karlsson, M. G., Zhou, M., Rogerson, T., Balaji, J., … Silva, A. J. (2009). CREB regulates excitability and the allocation of memory to subsets of neurons in the amyg-

dala. *Nature Neuroscience, 12*(11), 1438–1443.