Specific Aims

Understanding how experiences separated by time are associated together in memory is crucial since memories for individual events can only be understood in the framework of one's collective experience. Recently it was demonstrated in rodents that two neutral contexts experienced close in time can be bound in memory, a phenomenon termed memory linking. The goal of this proposal is to understand the ethological function of memory linking and how memory linking is supported by hippocampal ensembles.

It can be speculated that memory linking provides animals with the ability to predict environmentally relevant events, such as danger. For example, linking an aversive experience with an earlier experience may help the animal avoid future aversive events. If true, the temporal window of retrospective memory linking, where an animal links an aversive experience with a previously experienced neutral stimulus, should be greater than those for prospective memory linking, where an animal links an aversive experience with a neutral event experienced later in time. Notably, this ethological prediction differs dramatically from the prediction of the excitability hypothesis for memory linking, a previously proposed neurobiological explanation of memory linking. The excitability hypothesis states that the neurons encoding a memory have a transient increase in excitability, making them more likely to be incorporated into the ensemble of a memory formed shortly thereafter. Since negative valence increases neuron excitability, neurons engaged in a negative memory presumably sustain an elevated level of excitability for longer than those in a neutral memory. Thus, for a negative memory, the excitability hypothesis predicts a longer temporal window for prospective memory linking. The first aim of this proposal is to study the temporal window regarding the direction of memory linking. In line with recently acquired preliminary data, and we hypothesize that retrospective memory linking has a longer temporal window.

The second goal of this proposal is to study how the individual ensemble dynamics of two contextual memories contribute to the emergence of shared neuronal ensemble that supports memory linking. We have preliminary results suggesting that an increase in ensemble overlap for the linked contexts emerges during testing, but not during initial encoding. Thus, an intuitive hypothesis is that the representation of one of the contexts changes after encoding, so that it is more similar to the other context. We will investigate the reactivation rate for each context, which is the amount of overlap between the ensembles during retrieval and encoding. We hypothesize that one of the two ensembles will become more similar to the other. Moreover, we will assess whether activation patterns during initial encoding predict which cells are recruited into the shared neural ensemble for two linked contexts.

Lastly, calcium imaging in behaving animals using miniature microscopes is an essential tool for proposed studies. However, one of the challenges facing this approach is the analysis of imaging data. A non–negative matrix factorization (CNMF) algorithm has been shown to perform well on extracting calcium traces from raw video. However, lack of a user–friendly interface, batch–processing capability, and visualization tools of the results limits its popularity. Thus, the third goal of this proposal is to develop an analysis pipeline that provides an interface to the CNMF algorithm enabling the user to easily batch–process raw imaging data and visualize the results to quickly assess the quality of extraction. Such a tool will be very useful to the scientific community that employs calcium imaging with miniature microscopes.

Aim 1: Study the temporal window of prospective and retrospective memory linking. Test the hypothesis that temporal window for retrospective memory linking is longer compared to prospective memory linking.

Aim 2: Study how ensemble dynamic contribute to memory linking. Test whether ensemble of one context become more similar to the other linked context after encoding. Test whether different ensembles during encoding have equal contribution to the overlap between ensembles of linked context during retrieval. Test whether activity level of neurons in one session could predict whether they are active in future sessions.

Aim 3: Develop analysis pipeline for calcium imaging with miniature microscope.

A. Scientific Premise

It has been found in mice that two contextual memories for events that happened close in time can be linked together. This was evidenced by the ensembles of the two memories sharing a larger proportion of overlapping neurons in the hippocampus, and fear conditioning in one of the contexts resulting in freezing behavior in the other.[1] However, the temporal window for memory linking, as well as the neuronal basis of this phenomenon remains under–studied.

B. Background and Significance

Understanding how distinct memories can be related together is essential to understanding episodic memory, since memories for individual events can only be understood in the framework of one's collective experience. It is generally believed that different memories are encoded with different populations of neurons, which are reactivated during memory retrieval.[2] Such a specific population of neurons are termed the ensemble of said memory. Indeed, it has been found that artificially stimulating an ensemble is sufficient to drive the retrieval of the memory and elicit the associated behavioral response (*e.g.* freezing).[3] Thus, it can be speculated that two distinct memories can be associated together by sharing a proportion of neurons that is common to both ensembles, so that activation of one ensemble might trigger the activation of the other as well, potentially through a pattern completion mechanism. Consistent with this idea, it has been found that two types of aversive memories can be associated together by repeatedly presenting the stimuli that triggers each of them at the same time.[4] Importantly, the overlap between amygdala ensembles encoding the two aversive memories is larger than those in a control group that did not undergo the paired presentation of stimuli. Moreover, artificially suppressing the activity of the amygdala neuron population shared between the two ensembles disrupted the behavioral association between the two aversive memories, without affecting the independent recall of each of them.[4] Besides explicit pairing of two memories by triggering the recall of them at the same time, it has also been found that two fear memories can be associated together if they are encoded close in time. Specifically, two auditory fear conditioning memories can be associated together if they are encoded within 6 hours, but not when they are separated by 24 hours. Consistent with previous findings, the amygdala ensembles of the two memories that were associated together had a higher proportion of overlap.[5] It has been hypothesized that such a time-dependent association of different memories is mediated by an excitability-based memory allocation mechanism: Neurons encoding an earlier memory sustain an elevated level of excitability for certain period of time, thus biasing the allocation of a later memory towards the same population of neurons, resulting in an increase in overlap of the ensembles of the two memories.[6–8] Indeed, it has been found that artificially manipulating the amygdala neurons encoding one memory can bidirectionally bias the allocation of cells to a second memory and either rescue or disrupt the association between the two memories.[5] Taken together, these studies suggest that aversive memories can be associated together in a time-dependent manner. It was unknown, however, whether such association can occur naturally with memories that have neutral emotional valence, which is key to the question of how episodes of memories that does not have negative emotional valence can be linked together.

Recently, it was found in the rodent hippocampus that the neuronal ensembles of two neutral contextual memories separated by 5 hours have greater overlap than those separated by 2 days or 7 days. Moreover, subsequent fear conditioning in the latter of the two contexts initially experienced promotes elevated freezing levels in both that context and the linked context. Such results suggest a linking of two temporally distinct memories through overlapping neuronal ensembles. This phenomenon has been termed memory linking.[1]

How might neutral contextual memories experienced closer in time come to share greater overlapping hippocampal ensemble? It has been found that the representation of a familiar environment in the rodent hippocampus drifts across time, such that the similarity between ensembles for the same environment at different times decays as a function of temporal distance.[9,10] Such drift may reflect spontaneous turn–over of neural ensembles, which, when combined with the memory allocation hypothesis, could serve as a potential explanation of the phenomenon of memory linking. If two memories happen close in time so that when the second memory happens, neurons encoding the first memory still have elevated excitability, the two memories will have larger overlap in ensembles since the allocation of the second memory is biased towards the same neurons encoding the first memory. If, however, the two memories are separated by a large time interval, so that the excitability of neurons encoding the first memory return to baseline before the second memory happens, the two memories will have chance level overlap in ensembles due to spontaneous turn–over of ensembles over time. Taken together,

these results suggest that neutral memories can be linked in a time-dependent manner, possibly mediated by a spontaneous drift of ensembles and a memory allocation mechanism.

However, the temporal window of memory linking, which is the maximum time interval within which two memories can be linked together, remains under–studied. This question is important since it is closely related to the ethological implication of memory linking — It can be speculated that memory linking may play a role in learning the relationships between memory episodes, and the temporal window of memory linking reflects how much information is gathered to form relational memories for a specific episode. One of the most important relationship between memory episodes is causal relationship, where linking an aversive experience with an earlier experience may help the animal learn the causal relationship between events and avoid future aversive events. An important example is conditioned taste aversion (CTA), where the animals learn to associate sickness with the consumption of food several hours earlier, a period much longer than most stimulus–food associations.[11] In addition to highlighting the important biological function of linking events disparate in time, this example also highlights the predictive nature such associations can serve — CTA specifically ties sickness with stimuli experienced prior to sickness. Thus, retrospective memory linking, where an animal links an aversive experience with a the memory of a previously experienced neutral context, should be more important than prospective memory linking, where an animal links an aversive experience with a neutral contextual memory that happens later in time. Specifically, the temporal window of memory linking should be longer for retrospective memory linking than for prospective memory linking, since it helps the animal to gather more information about the potential cause of the aversive event. However, the allocation hypothesis would predict the opposite. Since it has been shown that memories with negative valence increase neuron excitability,[5,12] it is expected that neurons engaged in a negative memory sustain an elevated level of excitability longer than those in a neutral memory. Thus, for a negative memory, the allocation hypothesis would predict a longer temporal window for prospective memory linking compared to retrospective memory linking. In order to study the temporal window of prospective and retrospective memory linking, we have carried out preliminary behavior studies. The results suggest a longer temporal window for retrospective memory linking than for prospective memory linking. Thus, we hypothesize that retrospective memory linking has a longer temporal window than prospective memory linking.

Moreover, it is unclear how memory linking happens. As mentioned before, we have ethological speculations and preliminary data that cannot be fully addressed by the memory allocation hypothesis. We will focus on studying the mechanism of memory linking from a neural coding perspective. Specifically, we will focus on studying the neural correlates and ensemble dynamics of memory linking. Firstly, we have preliminary data suggesting that the increase in overlap of hippocampal ensembles emerges after the encoding of the two memories, which is contrary to what the allocation hypothesis would predict. To confirm this observation, we will compare the ensembles for each memory during encoding and retrieval to see whether there is a change after encoding. We will also compare the ensembles of one memory during encoding to the other during retrieval to see whether memory ensembles are becoming more similar to each other. Secondly, one of the most important neural correlates of memory linking is the ensemble overlap between the linked memories, since it has been shown that manipulating this overlapping population of neurons directly and specifically affects memory linking.[4] However, it is unclear how each memory contribute to the overlap. Thus, we will test whether the neurons engaged in the encoding sessions of each memory contribute equally to the overlap of ensembles during retrieval. Lastly, the allocation hypothesis predict that the activity level of a neuron during the encoding of the first memory should determine the likelihood of the said neuron being allocated to the second memory. However, this correlation has never been explicitly tested in previous memory linking studies. Thus, we will test whether the mean activity level during the encoding of the first memory differs for the neurons that are allocated to the second memories from those that are not. Taken together, these studies will give us more insight into how memory linking happens on a neural ensemble level.

Lastly, calcium imaging with miniature microscope in behaving animals is an important tool to study the neuronal dynamics in memory linking studies, due to its capability to track large populations of neurons across long periods of time. However, one of the challenges facing this technique is the analysis of imaging data. Specifically, extracting calcium traces of individual neurons from raw video is a difficult problem. Several approaches have been developed to address this difficulty. Most notably, a variant of the constrained non–negative matrix factorization (CNMF) algorithm has been shown to work well on calcium imaging data with miniature microscopes.[13,14] However, in practice, the accessibility of such approaches has been limited to the general scientific community that utilize miniature microscope for several reasons: a) The effect of different parameters in the mathematical model and the complications introduced by performance optimization of the code can be hard to comprehend

for researchers without a mathematical or programming background. b) Lack of a pre–processing pipeline that specifically addresses various artifacts usually observed in miniature microscope data requires the user to either rely on other software to conduct pre–processing steps or modify the existing code to extend its functionality. c) Lack of visualization tools at each step of analysis makes it hard for the researchers to check the quality of result. This is a significant issue especially since there are many detailed parameters that may have a huge effect on the final output of the CNMF algorithm. d) The lack of a batch-processing pipeline, as well as an easy way to visualize the results across multiple data–sets, limits the applicability of the algorithm to the analysis of large scale experiments. Thus, to address these difficulties, the final goal of the proposal is to develop an analysis pipeline that provides a user–friendly interface to the CNMF algorithm, and provides visualization tools as well as batch-processing capability, integrated into a single analysis package.

C. Preliminary Studies

Retrospective memory linking window is longer than prospective linking window for a negative memory

In the preliminary study shown in Figure 1, animals were put into two distinct contexts and separated into two major groups for which the temporal order of the two contexts during encoding is opposite. In the “prospective” group, animals received a delayed shock in context B first, and then explore context A, whereas in the “retrospective” group, animals explored context A first, and then get a delayed shock in context B. Both groups were tested in context A again 2 days after they are exposed to both context A and B. Within each major group, the animals were further divided into subgroups in which exposure to A and B were separated by varying intervals of time. Thus, elevated freezing level in the context A, where no shock ever occurred for both groups, indicates a transfer of fear and a linking of the two contexts. Furthermore, comparison of freezing levels in context A between subgroups with various time intervals can reveal the difference of temporal windows of memory linking across “prospective” and “retrospective” groups. For all groups, the exploration and testing session lasted for 5 minutes, a foot shock of 0.5 mA was delivered at the beginning of the second minute. The various time intervals are 5 hours, 1 day, 2 days and 7 days for both groups.

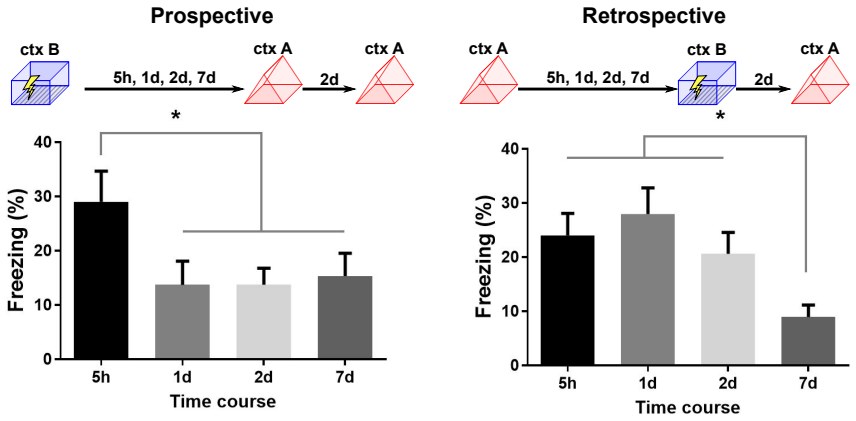

Figure 1: Retrospective memory linking window is longer than prospective linking window for a negative memory.

In the “prospective” linking group, we observed significant higher freezing level in context A with 5 hours interval relative to those with 7 day interval, but not with either 1 day or 2 days interval. This suggest that the fear memories were able to transfer forward to a neutral context 5 hours in the future. On the other hand, in the “retrospective” linking group, we observed higher freezing level with 5 hours, 1 day and 2 days interval compared to those with 7 days interval. This suggests that fear memory was able to transfer backward up to 2 days. Taken together, these results suggest an extended temporal window of retrospective memory linking compared to prospective memory linking.

The difference in overlap of ensembles across groups emerge after encoding

In the preliminary study shown in Figure 2, the effect of negative valence on the temporal window of retrospective memory linking was studied. Briefly, animals were put into context A and B in that order and were divided into two major groups — “neutral” and “aversive”. The difference between the two major groups is that in the “aversive” group, the animals received a delayed shock during the encoding session of context B, thus making context B an aversive memory from the beginning. In contrast, for the “neutral” group, the emotional valence of context B was initially neutral, and the animals received an immediate shock 2 days after the initial exploration of context B. Similar to the preliminary experiments described earlier, animals within each major group were further divided into sub groups according to various time intervals between initial exposure of context A and B, while all

of the animals were put back to context A for testing and the comparison of freezing levels in A between sub groups could reveal the difference in temporal windows of memory linking across “neutral” and “aversive” groups. We observed that consistent with published works, animals in the “neutral” group can only link the two memories across 5 hours. However, animals in the “aversive” group can link the two memories across 2 days, suggesting an extension of temporal window of memory linking due to emotional valence of the initial context memory. Thus, we chose a temporal interval of 2 days for following calcium imaging studies, since there was a significant difference in behavior across the two groups with a 2 day time interval.

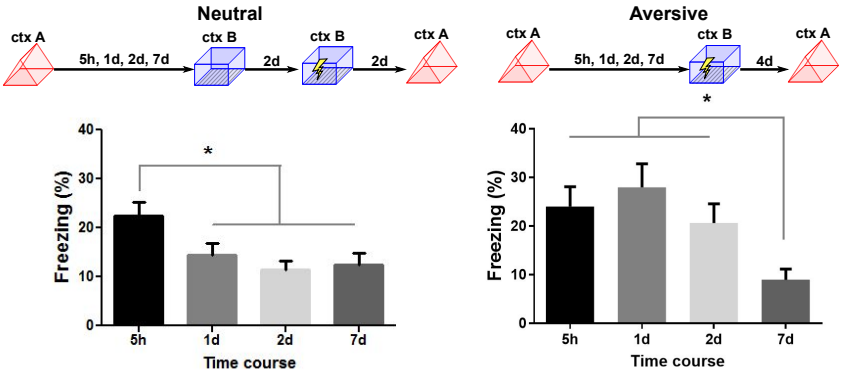

Figure 2: Negative valence increased temporal window of memory linking.

In the accompanying imaging study shown in Figure 3, the animals were divided into two groups and undergo the same paradigm as described above. The difference is that the time interval between context A and B is set to 2 days for all animals, and that animals are tested in a novel context C in addition to context A. Calcium imaging with miniature microscope was carried out in all sessions, so that the neural ensemble of each session can be obtained and comparisons of ensembles can be carried out within subject. Specifically, we looked at the overlap between two ensembles, which is the number of neurons that were active in both sessions as a proportion of the total number of neurons that were active in either sessions. Consistent with behavioral results, during retrieval, we observed a significant increase in the overlap between ensembles of context A and B in the “aversive” group, but not in the “neutral” group. Moreover, no difference was observed between the two groups for the ensemble overlaps involving the novel context C, *i.e.* the ensemble overlaps between context A and C or between context B and C, confirming that the increase in ensemble overlap between context A and B is a specific neural correlate for memory linking. Surprisingly, no difference in ensemble overlap between context A and B was observed across the two groups in encoding sessions. Such a result suggests that the increase in ensemble overlap, which is hypothesized to drive memory linking, emerges after encoding, which is contrary to what would be predicted by the memory allocation hypothesis.

Minian: a python analysis pipeline for calcium imaging data

To address the difficulties raised during the analysis of calcium imaging data with miniature microscopes, we have developed an open-source analysis pipeline that allows the user to easily run through code, take notes and visualize the results at the same time. Specifically, the pipeline incorporates pre–processing steps that take raw videos as inputs, and corrects various artifacts that are usually seen with miniature microscopes. For instance, the motion correction step for miniature microscope is usually hindered by a “dark ring” around the field–of–view caused by vignetting, “bright spots” caused by dust on the lens or sensor, and a noise pattern that is stable across time.

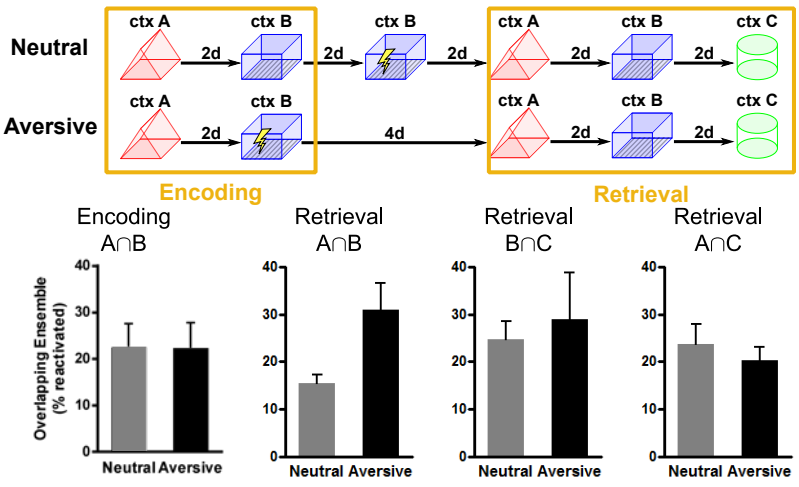

Figure 3: Increase in ensemble overlap emerges after encoding.

We apply various interpolation and smoothing steps to specifically remove such artifacts before motion correction, and the result of motion correction is improved. The results from each step, including motion correction, can be visualized as dynamically–generated animations (*i.e.* without saving intermediate videos to hard drives). Thus the user can easily visualize the effect of pre–processing steps and decide whether each step is necessary for the specific data. Regarding the result of the CNMF algorithm, there are two major concerns about the quality of the extraction of calcium traces: a) the algorithm usually picks up “units” around the border of the field–of–view

that are actually noise. b) the balance between “demixing”, where overlapping neurons are separated and treated as individual neurons, and “merging”, where spatially scattered fluorescent signals coming from the same neuron are treated as single source, can be very sensitive to parameter tweaking, and it is often the case that signals coming from one neuron are falsely treated as separated neurons, or vice versa. Thus, the pipeline also provides visualization tools to address these issues. Specifically, the tools provide an overlaid contour plot of the spatial footprints of all identified “units” within a session, thus the shape of each “unit” can be easily inspected. Moreover, the user can directly select “units” of interest from the overlaid contour plots, and the calcium traces of selected “units” can be visualized in a manner that is similar to the plot of spike trains of different neurons. Thus, the user can easily inspect whether “units” that are overlapping or close to each other have correlated calcium traces that suggest a single source of signal. Furthermore, there is an option to manually accept or reject “units” within the visualization tool, so that the user can modify the results and take out false “units” without leaving the visualization. All these features help the user to detect and correct the potential pitfalls of the calcium signal extraction process. Finally, the whole pipeline is provided as a jupyter notebook, which allows the user to easily take and share text notes, edit and execute code, and dynamically visualize results all in an integrated interface.

D. Research Design and Methods

Aim 1: Study the temporal window of prospective and retrospective memory linking.

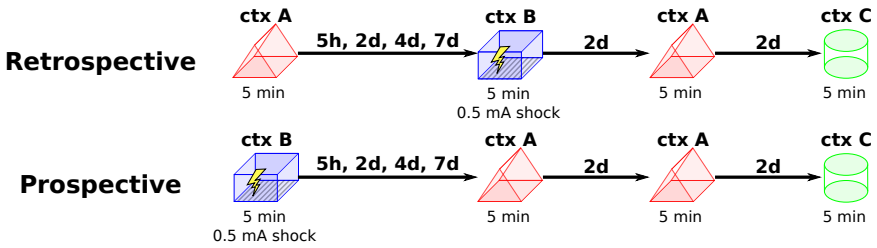

Figure 4: behavior experimental design.

To test hypothesis that retrospective memory linking has a longer temporal window than prospective memory linking, we will carry out behavior studies as shown in Figure 4. Specifically, animals will be put into two distinct contexts and separated into two major groups where the temporal order of the two contexts during encoding is opposite. In the “prospective” group, animals will receive a delayed shock in context B first, and then explore context A, whereas in “retrospective” group, animals will explore context A first, and then get a delayed shock in context B. For both groups, the exploration sessions will last for 5 minutes, and a foot shock of 0.5 mA will be delivered at second minute during the encoding session of context B. After encoding, both groups will be tested in context A and a novel context C in that order 2 days after they are exposed to both context A and B. The time interval between the two testing session will be 2 days. During testing, animals’ behavior will be recorded, and freezing level will be assessed from the videos using standard software (Med Associates, Inc.). Within each major group, the animals will be further divided into subgroups, and will be equally assigned to either 5 hours, 1 day, 2 days or 7 days subgroup, which indicates the time interval between the exposure of context A and B. Then, within each subgroup, animals’ freezing level in context A during testing can be compared with those in the novel context C. An elevated freezing level in context A, where no shock ever occurred for all groups, relative to the novel context C, indicates a transfer of fear from context B to context A. Then within each major group, we can estimate the temporal window of memory linking by looking at whether context A and B are linked together with each temporal intervals. Finally, the temporal window for retrospective and prospective memory linking can be compared between the two major groups. We expect to see retrospective memory linking across 2 days, while the temporal window of prospective memory linking will be shorter than 1 day.

The presented experimental design has 8 unique groups in total. We plan to have 12 animals assigned to each group, resulting in a total of 96 animals for the presented study. The experiment is expected to be finished within 6 months.

One concern of the presented experimental design is the order effect of testing sequence, where testing in one of the context will affect freezing level in the other context that is tested later. To control for this, we will switch the testing order of context A and context C in selected groups. If order effect are observed, we can adopt an alternative “parallel” testing design, where animals within each group can be further divided into two groups, with half of the animals being tested in context A only, and the other half being tested in context C only.

Aim 2: Study how ensemble dynamic contribute to memory linking.

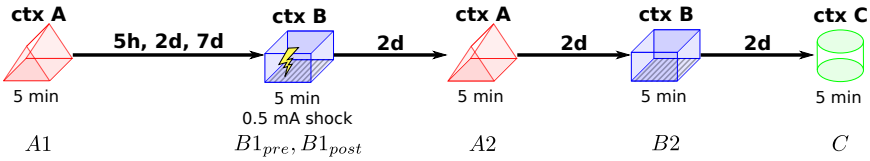

Figure 5: imaging experiment design.

To study how ensemble dynamics contribute to memory linking, we will carry out calcium imaging studies shown in Figure 5. The design is similar to the "retrospective" group in the behavior experiments proposed in Specific Aim 1. Briefly, the animals will be placed in context A to explore, then after either 5 hours, 2 days or 7 days, the animals will be placed in context B, where they receive a delayed shock at second minute. Then, two days later, the animals are put back in context A, context C, and context B in that order for testing. We chose this design since in preliminary studies, we observed an emergence of ensemble overlap between context A and B after encoding, with a 2 day time interval between encoding of context A and B. At the same time, we expect to see memory linking when encoding of context A and B are separated by 5 hours, but not when they are separated by 7 days. Thus, the 5 hours and 7 days groups serve as positive and negative controls, respectively.

Before the experiment, all animals will be subjected to surgery procedures where virus expressing GCaMP6 is injected unilaterally into the dorsal CA1 region of hippocampus and grin lenses are implanted directly above the injection site. A base plate will also be attached to the skull to secure the grin lens. During behavioral training and testing, miniature microscopes will be attached to the base plate and calcium signals will be recorded in all sessions. The recorded video will be processed with an analysis pipeline described in Specific Aim 3, and calcium traces for individual neurons can be extracted with CNMF algorithm. Thus the neural dynamics during all sessions can be collected and analyzed.

$$
\begin{aligned}
R_{A1} &= \frac{N\{A1 \cap A2\}}{N\{A1\}} \\
R_{B1pre} &= \frac{N\{B1_{pre} \cap B2\}}{N\{B1_{pre}\}} \\
R_{B1post} &= \frac{N\{B1_{post} \cap B2\}}{N\{B1_{post}\}}
\end{aligned} \quad (1)
$$

For simplicity, we use uppercase letters combined with numbers to denote the population of neurons that were active during a recording session as shown in the bottom of Figure 5: we use $A1$ and $A2$ to denote neurons that are active during the encoding and retrieval session of context A, respectively. We use $B1_{pre}$ and $B1_{post}$ to denote cells that active before and after the delivery of the shock during the encoding session of context B. In addition, we use $B2$ to denote active neurons in the retrieval session of context B, and lastly we use C to denote the active neurons encoding context C. Additionally, we use set operation conventions to denote the specific population of neurons that we are interested in. Specifically, we use set intersection "$\cap$"to denote the population of neurons that are active in both sessions. For example, $A1 \cap A2$ denotes the cells that are active in both the encoding and the retrieval session of context A. Moreover, we use set difference "$\setminus$"to denote the population of neurons that are active in the first session but not the second session. For example, $A1 \setminus A2$ denotes the cells that are active only during the encoding session of context A, but not reactivated during the retrieval session of context A. Finally, we use "$N\{\ldots\}$" to denote the number of cells in the population that we are interested in. For example, $N\{A1\}$ denotes the total count of cells that are active during the encoding of context A.

To test **whether representation of one context become more similar to the other after encoding**, we will first look at the reactivation rate of the neurons initially encoding a context during the retrieval of that context. Specifically, we will calculate reactivation rate as shown in Equation 1, where R_{A1}, R_{B1pre}, R_{B1post} denote the reactivation rate of the ensembles for context A, pre-shock phase for context B, and post-shock phase for context B, respectively. In addition, we will also look at whether one context becomes more similar to the other after encoding. To test this, we can calculate a reactivation rate of the neurons initially encoding one context during the retrieval of the other context. Specifically, we will calculate the rate as shown in Equation 2. Conceptually, R_{BA} indicate how close is the representation of context B during retrieval to the initial representation of context A. Similarly, R_{ABpre} and R_{ABpost}indicate how close is the representation of context A during retrieval to the initial pre-shock or post-shock representation of context B, respectively. If there is a drift of the representation of one context towards the other after encoding, we expect to see a decrease of reactivation rate of that context between encoding and retrieval, as well as a higher overlap between the retrieval of that context and the encoding of the other context.

$$
\begin{aligned}
R_{BA} &= \frac{N\{A1 \cap B2\}}{N\{A1\}} \\
R_{ABpre} &= \frac{N\{B1_{pre} \cap A2\}}{N\{B1_{pre}\}} \\
R_{ABpost} &= \frac{N\{B1_{post} \cap A2\}}{N\{B1_{post}\}}
\end{aligned} \quad (2)
$$

To study **how different ensembles during encoding may contribute to the overlap during the retrieval**,

we will calculate the proportion of cells during encoding that also serve as the overlapping cells during retrieval, as shown in Equation 3. X is one of the three ensembles during encoding, that is: $A1$, $B1_{pre}$, or $B1_{post}$. The term $(A2 \cap B2) \setminus (A2 \cap B2 \cap C)$ give us the population of neurons that are active specifically in both the retrieval of context A and B, but not active during all three retrieval sessions of context A, B and C. The intersection between this population and the ensembles of one of the encoding sessions will then estimate the contribution of neurons from one of the encoding sessions to this specific population that may drive memory linking. Thus, C_{A1}, C_{B1pre} and C_{B1post} will indicate the contributions to the overlap during retrieval from the ensembles encoding context A, pre-shock phase of context B, and post-shock phase of context B respectively. We will then compare these three rate to see whether different ensembles during encoding contribute equally to the population of neurons that may drive memory linking.

$$C_X = \frac{N\{X \cap (A2 \cap B2) \setminus (A2 \cap B2 \cap C)\}}{N\{X\}} \quad (3)$$

Lastly, in line with the allocation hypothesis, we expect **the activity level of a neuron should affect the likelihood of that neuron being reactivated during future sessions**. Specifically, for each neuron, the activity level of that neuron can be summarized by the area-under-the-curve of the calcium trace of that neuron. By taking a z-score of such activity level across all neurons in a recording session, we can obtained a centered activity level for each neuron that's also normalized to the overall activity level of all neurons in that session. We can denote such centered and normalized activity level in session X as $A_X\{\ldots\}$, where $\ldots$ indicate a specific population we are interested in. Then, for any given two recording sessions X and Y, where X happens before Y, four quantities will be calculated and compared as shown in Equation 4. $Activity_X^{common}$ indicate the activity level in X of those cells that were reactivated in Y, whereas $Activity_X^{alone}$ indicate the activity levels in X of the cells that were only activated during X, but not Y. We will then compare the mean of $Activity_X^{common}$ and $Activity_X^{alone}$. According to allocation hypothesis, we expect to see higher $Activity_X^{common}$ than $Activity_X^{alone}$, since the activity level in X should affect whether they are reactivated in later sessions. On the other hand, we expect no difference between $Activity_Y^{common}$ and $Activity_Y^{alone}$, since the activity level of cells in the later session should not affect whether they are co-activated in both sessions.

$$\begin{aligned} Activity_X^{common} &= A_X\{X \cap Y\} \\ Activity_X^{alone} &= A_X\{X \setminus (X \cap Y)\} \\ Activity_Y^{common} &= A_Y\{X \cap Y\} \\ Activity_Y^{alone} &= A_Y\{Y \setminus (X \cap Y)\} \end{aligned} \quad (4)$$

The presented experiment has 3 groups in total. We plan to assign 8 animals for each group, resulting in a total of 24 animals. The surgeries, experiments and analysis are expected to be finished in 18 months.

Besides the analysis proposed above, the rich dataset collected from this experiment provides opportunities for other types of analysis. For example, it might be the case that each encoding session contributes equally to the ensemble overlap between the two linked memories. In other words, the encoding session in which a neuron was active cannot predict the likelihood of the said neuron to be recruited into the overlapping population of ensembles. Instead, it is possible that the overlapping population is biased towards those neurons that were encoding an internal state that was common to both memories. For instance, neurons correlated with shock delivery or freezing might be preferentially recruited as the overlapping neurons between the two ensembles. To investigate this possibility, we can correlate the temporal dynamics of the neurons in the overlapping population to some external variables such as shock onset or freezing episodes, and see whether neurons encoding this information make up a high proportion of the overlapping population. Furthermore, the neurons in the overlapping population of two ensembles might encode information that can not be correlate with an explicit variable. For instance, they might encode an abstract "temporal context". In such cases, dimension reduction tools that can uncover the structure of neuronal dynamics in an unsupervised way might help us understand how memory linking happens in general.

Aim 3: Develop analysis pipeline for calcium imaging with miniature microscope.

The existing analysis pipeline described in the preliminary results still lacks several important features — the capability to be easily run in batch mode and the ability to easily visualize the results from multiple recording sessions for a large scale dataset. Several aspects are important to consider when developing this feature. a) The metadata, such as session i.d., animal and experiment need to be properly stored and associated with the result of CNMF in order to properly organize and index the dataset as a whole. The Python package "xarray" provides a solution by supporting a labeled N–dimensional data structure that incorporates data with metadata (labels) in an intuitive manner. b) The size of a large dataset usually cannot fit into the RAM space of a personal computer. Thus, to visualize and analyze the result of CNMF for large–scale datasets, a "lazy loading" feature is required

where the data are only loaded into the RAM when needed (*i.e* when the actual computation happens), and will be cleared out of the RAM to make space for other data once the computation is finished. The Python package “dask” provide this functionality and support an interface to “xarray” out–of–box. c) The visualization of a large dataset can be expensive to generate in terms of both time and disk space. Thus, the capability to generate visualization dynamically is important. The Python package “holoviews” provides this functionality and enable easy visualization of large datasets. It also supports an interface to “xarray”. We will use the package mentioned above to further develop the analysis pipeline for calcium imaging with miniature microscopes. Such a tool will be very valuable to the scientific community.

References

[1] D. J. Cai, D. Aharoni, T. Shuman, J. Shobe, J. Biane, W. Song, B. Wei, M. Veshkini, M. La-Vu, J. Lou, S. E. Flores, I. Kim, Y. Sano, M. Zhou, K. Baumgaertel, A. Lavi, M. Kamata, M. Tuszynski, M. Mayford, P. Golshani, and A. J. Silva. A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, 534(7605):115–118, June 2016.

[2] K. Tayler, K. Tanaka, L. Reijmers, and B. Wiltgen. Reactivation of Neural Ensembles during the Retrieval of Recent and Remote Memory. *Current Biology*, 23(2):99–106, January 2013.

[3] S. Ramirez, X. Liu, P. P.-A. Lin, J. Suh, M. Pignatelli, R. L. Redondo, T. J. Ryan, and S. Tonegawa. Creating a false memory in the hippocampus. *Science*, 341(July):387–392, July 2013.

[4] J. Yokose, R. Okubo-Suzuki, M. Nomoto, N. Ohkawa, H. Nishizono, A. Suzuki, M. Matsuo, S. Tsujimura, Y. Takahashi, M. Nagase, A. M. Watabe, M. Sasahara, F. Kato, and K. Inokuchi. Overlapping memory trace indispensable for linking, but not recalling, individual memories. *Science*, 355(6323):398–403, January 2017. PMID: 28126819.

[5] A. J. Rashid, C. Yan, V. Mercaldo, H.-L. L. Hsiang, S. Park, C. J. Cole, A. D. Cristofaro, J. Yu, C. Ramakrishnan, S. Y. Lee, K. Deisseroth, P. W. Frankland, and S. A. Josselyn. Competition between engrams influences fear memory formation and recall. *Science*, 353(6297):383–387, July 2016. PMID: 27463673.

[6] J. Lisman, K. Cooper, M. Sehgal, and A. J. Silva. Memory formation depends on both synapse-specific modifications of synaptic strength and cell-specific increases in excitability. *Nature Neuroscience*, page 1, February 2018.

[7] A. Yiu, V. Mercaldo, C. Yan, B. Richards, A. Rashid, H.-L. Hsiang, J. Pressey, V. Mahadevan, M. Tran, S. Kushner, M. Woodin, P. Frankland, and S. Josselyn. Neurons Are Recruited to a Memory Trace Based on Relative Neuronal Excitability Immediately before Training. *Neuron*, 83(3):722–735, August 2014.

[8] Y. Zhou, J. Won, M. G. Karlsson, M. Zhou, T. Rogerson, J. Balaji, R. Neve, P. Poirazi, and A. J. Silva. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. *Nature Neuroscience*, 12(11):1438–1443, November 2009.

[9] E. A. Mankin, F. T. Sparks, B. Slayyeh, R. J. Sutherland, S. Leutgeb, and J. K. Leutgeb. Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, 109(47):19462–19467, November 2012. PMID: 23132944.

[10] Y. Ziv, L. D. Burns, E. D. Cocker, E. O. Hamel, K. K. Ghosh, L. J. Kitch, A. E. Gamal, and M. J. Schnitzer. Long-term dynamics of CA1 hippocampal place codes. *Nature Neuroscience*, 16(3):264–266, March 2013.

[11] J. Garcia, D. J. Kimeldorf, and R. A. Koelling. Conditioned Aversion to Saccharin Resulting from Exposure to Gamma Radiation. *Science*, 122(3160):157–158, July 1955. PMID: 14396377.

[12] B. M. McKay, E. A. Matthews, F. A. Oliveira, and J. F. Disterhoft. Intrinsic Neuronal Excitability Is Reversibly Altered by a Single Experience in Fear Conditioning. *Journal of Neurophysiology*, 102(5):2763–2770, September 2009.

[13] E. A. Pnevmatikakis, D. Soudry, Y. Gao, T. A. Machado, J. Merel, D. Pfau, T. Reardon, Y. Mu, C. Lacefield, W. Yang, M. Ahrens, R. Bruno, T. M. Jessell, D. S. Peterka, R. Yuste, and L. Paninski. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron*, 89(2):285–299, January 2016. PMID: 26774160.

[14] P. Zhou, S. L. Resendez, G. D. Stuber, R. E. Kass, and L. Paninski. Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *arXiv:1605.07266 [q-bio, stat]*, May 2016.