

2023

A scale-invariant representation of temporal context: support and constraints from neural and behavioral data

<https://hdl.handle.net/2144/48503>

"Downloaded from OpenBU. Boston University's institutional repository."

BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**A SCALE-INVARIANT REPRESENTATION OF
TEMPORAL CONTEXT: SUPPORT AND
CONSTRAINTS FROM NEURAL AND BEHAVIORAL
DATA**

by

IAN M. BRIGHT

B.S., Indiana University of Pennsylvania, 2016

B.A., Indiana University of Pennsylvania, 2016

M.A., Boston University, 2019

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2023

© 2023 by
IAN M. BRIGHT
All rights reserved

Approved by

First Reader

Marc W. Howard, Ph.D.
Professor of Psychological and Brain Sciences

Second Reader

Michael E. Hasselmo, D.Phil.
William Fairfield Warren Distinguished Professor
Professor of Psychological and Brain Sciences

Third Reader

Sam Ling, Ph.D.
Associate Professor of Psychological and Brain Sciences

*I remember the time I knew what happiness was;
let the memory live again.*

-Cats (2019)

Acknowledgments

This was initially a very rough draft I wrote to make sure the formatting was correct during review. I never got around to making it sound better. You can read this, but it isn't very good so I'd rather you didn't. Your call though. I'll probably start by talking about how this is a journey I didn't take alone. I wouldn't be where I am now if it wasn't for a lot of other people, and its time they get some credit. Gonna talk about my parents here, and how they supported me, fostered my intellectual curiosity, and put up with me rarely doing my homework. Now I'm pivoting to close friends (non-academic) that kept me grounded in reality. Then I'll talk about undergrad, and how surrounding myself with a bunch of high achieving nerds turned out to be a good thing. Gotta shout out the honors college for paying for me to go on a vacation (to a conference), but in a way that sounds better than that. Without that trip I would never have made the jump to psychology. Then I'll talk about Dr. Pavloski giving some math major a start in psychology research.

Now I'll talk about my grad school classmates. In particular I have to thank Claudio and Michael. Then its time to talk about the lab. Everyone mattered to me but I gotta specifically mention a few people. Thanks to Inder for helping me understand continuous recognition, and all around welcoming me to the lab while simultaneously trying to wrap up his doctorate and plan a wedding. Thanks to Zoran for helping me make sense of the difference between " τ " vs " τ^* " and "big T " vs "little t ". I don't know if I'd understand the models like I do without him. Thanks to Nathan for helping me learn how to do neural data analysis and indulging my rants when I didn't feel like working. Thanks to Rui for pushing me to think about the stuff I'd never consider, always having thoughtful feedback on my work, and being a great all around friend.

Gonna thank Sam for having the thought to give my application to Marc as a

pivot to talking about Marc. Finally I have to try and sum up my thanks to Marc. Genuinely not sure where to start on that. For seeing something in a clueless kid (particularly when it came to memory) I'm still not sure I see. For showing incredible patience during my struggles in my first year. For handing me an incredible first neural dataset to analyze before I had the chance to ask for it. For showing even more patience during the pandemic. For giving me the tools, knowledge, and time required for me to grow into someone who is hopefully ready to take the next steps in my journey. I can only hope I'm half the mentor Marc is.

**A SCALE-INVARIANT REPRESENTATION OF
TEMPORAL CONTEXT: SUPPORT AND
CONSTRAINTS FROM NEURAL AND BEHAVIORAL
DATA**

IAN M. BRIGHT

Boston University, Graduate School of Arts and Sciences, 2023

Major Professor: Marc W. Howard, Ph.D.

Professor of Psychological and Brain Sciences

ABSTRACT

Episodic memory retrieval is believed to require recovering a slowly changing spatiotemporal context. Two findings supporting this view are the recency effect (recent events are more likely to be remembered) and the contiguity effect (after recollecting an event, things that occurred near the event are more likely to be remembered). Critically, contiguity and recency effects are similar across time scales (i.e., they are scale-invariant). As a result, a scale-invariant model of context was proposed which makes several empirical predictions about the brain and behavior. Neurally, this model predicts neurons that respond shortly after an event and exponentially return to baseline at various rates. Experiment 1 analyzed single-unit recordings from the entorhinal cortex of 2 male rhesus macaques as they freely viewed images. In support of this hypothesis, I report units that fired in response to image onset and then relaxed their firing with a spectrum of rates.

Behaviorally, this model suggests that the time required to initiate memory retrieval can increase logarithmically with recency. Experiment 2 analyzed results from

6 continuous recognition experiments, where 202 human adults judged on each trial if a stimulus is new or old without separate study and test phases. Response time distributions demonstrated that the time required to initiate memory retrieval increased as a logarithmic function of that memory's recency out to at least 8 minutes, supporting this behavioral hypothesis.

Finally, although there is a spectrum of decay rates, there must be a fastest decay rate—items presented at a rate faster than this fastest decay rate should disrupt the contiguity effect. Experiment 3 analyzed behavioral results from 330 human adults performing an immediate free recall task in which lists were presented at 2, 4, or 8 Hz. In immediate free recall, participants studied a list of items and began recall as soon as the list ended. Critically, for lists presented at 8 Hz, the contiguity effect was eliminated, consistent with the hypothesis that the contiguity effect does not occur at very short time scales. Together these results suggest that temporal context is scale-invariant and offer additional constraints to the form of this representation.

Contents

1	Introduction	1
1.1	Behavioral effects in memory	1
1.2	Retrieved context models	4
1.3	A scale-invariant representation of temporal context	6
1.4	Neural representations of time	7
1.5	Overview of chapters	9
2	A temporal record of the past with a spectrum of time constants in the monkey entorhinal cortex.	12
2.1	Introduction	12
2.2	Results	15
2.2.1	Temporal receptive field analysis revealed a population of temporal context cells in entorhinal cortex	18
2.2.2	Populations of entorhinal temporal context cells carry graded information about time	25
2.2.3	EC neurons conveyed information about image identity	29
2.3	Discussion	33
2.3.1	Relationship to findings from rodent and monkey EC	34
2.3.2	Exponentially-decaying neurons with a spectrum of time constants is the Laplace transform of time	36
2.4	Methods	38
2.4.1	Subjects, training, and surgery	38

2.4.2	Electrophysiology	39
2.4.3	Experimental design and behavioral task	40
2.4.4	Analysis of Neural Firing Fields	41
2.4.5	Linear Discriminant Analysis (LDA)	43
2.4.6	Stimulus sparsity analysis	45
2.4.7	Population vector analysis of stimulus specificity	45
3	The time to initiate retrieval of a memory depends on recency.	47
3.1	Methods	51
3.1.1	Experiment 1	51
3.1.2	Experiment 2	55
3.1.3	Experiment 3	56
3.1.4	Experiment 4	57
3.1.5	Experiment 5	59
3.1.6	Experiment 6	60
3.1.7	Analyses	61
3.2	Results	62
3.2.1	First Repetitions	62
3.2.2	Multiple Repetitions	69
3.3	General Discussion	73
4	Rapid presentation rate negatively impacts the contiguity effect in free recall.	77
4.1	Introduction	77
4.2	Methods	79
4.2.1	Participants	79
4.2.2	Procedure	80
4.2.3	Analysis	80

4.3	Results	81
4.3.1	As Presentation Rate Increases, Fewer Words are Recalled . .	81
4.3.2	Increasing Presentation Rates Increases the Primacy Effect and Decreases the Recency Effect	82
4.3.3	The Contiguity Effect Flattens At Higher Presentation Rates .	83
4.4	Discussion	86
5	Conclusions	88
5.1	Summary of dissertation	88
5.1.1	Chapter 2 results	88
5.1.2	Chapter 3 results	89
5.1.3	Chapter 4 results	89
5.2	Future directions	90
5.2.1	Temporal context cells for the past and future	90
5.2.2	Connections between episodic memory and sharp-wave ripples	91
5.2.3	Drift at multiple time scales in the human anterior temporal lobe	94
	References	95
	Curriculum Vitae	113

List of Tables

3.1	The major methodological differences between the six continuous recognition experiments. The experiments varied on the subject pool that participants were recruited from (Amazon M-Turk or Boston University Students), what was displayed between each trial (fixation cross or phase scrambled noise), the type of stimuli used (images or words), the instructions given to the participants (respond only to repeated items or indicate if the current item is new/old), and the number of times an item could be repeated (2 or 5).	53
3.2	Slope and intercept values from a logistic mixed effects regression of lag on hit rate for first repetition items. All tests were performed on the base 2 logarithm of lag. Bold z-scores indicate significance at the $p < 0.05$ level. Across all six experiments, hit rates decreased as a function of lag.	64
3.3	Slope and intercept values from a linear mixed effects regression of lag on median response times for first repetition items. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times increased as a function of lag.	66

3.4	Slope and intercept values from a linear mixed effects regression of lag on first decile response times for first repetition items. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times increased as a function of lag.	67
3.5	The slopes and intercepts of the linear mixed effects regression on Lag Modulation Factor as a function of Quantile across the six experiments for first repetition items. Bold t-scores indicate significance at the $p < 0.05$ level. The intercepts calculated from the lag modulation factor analysis were significant in all six experiments.	69
3.6	Slope and intercept values from a linear mixed effects regression of the lag between an item's second and third presentation (lag_2) and the lag between an item's first and second presentation (lag_1) and on median response times for items repeated a second time. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times increased as a function of lag_2 but did not systematically vary for lag_1	71
3.7	Slope and intercept values from a linear mixed effects regression of lag and number of repetitions on median response times. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times decreased with number of repetitions, but the effect of lag was the same regardless of number of repetitions.	72

List of Figures

1.1	A walk along alphabet hill. <i>Top:</i> An illustration of an individual walking along a hill with letters on it. At each moment, the thoughts of the individual, represented as thought bubbles, includes the recent past. This is their temporal context (e.g., A, B, C). <i>Bottom:</i> Jumping back in time. When asked to remember if they saw the letter ‘C’, the individual retrieves the entire context associated with ‘C’. Adapted from Folkerts et al., 2018.	2
1.2	Scale-invariant representations of temporal context. a. Peristimulus time histograms (PSTH) of hypothetical cells in the first layer F . Lighter lines correspond to cells with longer relaxation times s . b. PSTH of hypothetical cells in the second layer T . Lighter lines correspond to cells with later peak times $\tilde{\tau}^*$. c. Heatplot of a population of hypothetical cells in F , sorted by their corresponding relaxation time, where yellow corresponds to high firing and blue low firing. d. Same format as c but for hypothetical cells in T sorted by their corresponding peak time.	8

2.1 **a-b: Experimental procedures.** **a**, On each trial, the monkey freely viewed an image. After the monkey viewed an image for 5 seconds, the image disappeared. Following every trial, the monkey performed gaze calibration trials and received a reward. Images were presented twice during an experimental session. **b**, Estimated recording channel location in the entorhinal cortex for one recording session is shown in red on a coronal MRI. **c-e: Two hypotheses for neural representations of time.** **c-d**, Heat plot (top) and tuning curves (bottom) for two hypotheses for how a time interval of image viewing could be represented. In the heat plots, cooler colors correspond to low activity while warmer colors correspond to higher activity. **c**, Hypothetical activity for sequentially activated time cells, like those observed in the hippocampus. Because the time of peak response across neurons covaries with the spread of the firing field, neurons with later firing fields display wider firing fields. **d**, Hypothetical activity for monotonically decaying temporal context cells, like those observed in rodent EC. Neurons in this population reach their peak at about the same time. However, different neurons decay at different rates. **e**, Properties of neurons representing time passage by the hypothetical cells in **c** and **d**. A population of time cells (red) should respond at different points in a time interval, and show a robust correlation between response latency and reaction time. Conversely, a population of temporal context cells (gray) should respond in a more restricted time range shortly after the start of a time interval, and show no correlation between response latencies and relaxation times. baseline.

2.2 **Modeling neuron responses using an ex-Gaussian distribution.** **a**, The ex-Gaussian model of a neuron’s response (bottom) is formed by convolving a Gaussian distribution (with parameters μ and σ) with an exponential distribution (with parameter τ) (top). **b**, In the ex-Gaussian model, an increase in μ shifts the distribution to the right, an increase in σ widens the central peak of the distribution, and an increase in τ lengthens its decay rate. The gray and red arrows correspond to the parameter changes expected in the process of fitting the response of temporal context cells and sequentially activated time cells respectively.

2.3	Temporal context cells in monkey entorhinal cortex respond shortly after image presentation, then return to baseline at a spectrum of rates.	a, Three representative cells that responded to image onset and decayed at different rates (Figure 2.4 shows more examples). Each pair of plots indicate the activity of one neuron relative to image onset in a raster plot (top) and PSTH (bottom). In each raster, a tick mark indicates when the neuron fired an action potential. In the PSTH, the solid black line indicates the smoothed firing rate, the pink line indicates the model estimate of firing rate, and the estimated baseline firing rate is indicated by the black dotted line. Relaxation Time refers to the duration between response peak and when the neurons returned 63% of the way to baseline. Pearson's r is the correlation between the fits for even and odd trials.	b, A heatplot of the normalized firing rate of 109 temporal context cells, relative to image onset, sorted by their Relaxation Time. The color scheme is the same as in Figure 2.1c, d. The majority of neurons responded within 1 s of image onset, relaxing back to baseline with a spectrum of decay rates.	c, A scatter plot of the joint distribution of each neuron's Response Latency and Relaxation Time. Response Latencies did not span the entire 5 s, unlike Relaxation Time, and a neuron's Response Latency and Relaxation Time were not correlated. Figure 2.5 shows the marginal distributions of these parameters.	19
2.4	Additional examples of temporal context cells.	Format is as in Figure 2.3a.			21

2.5 **The distribution of individual parameters is consistent with temporal context cells.** **a**, Neurons started responding (Response Latency) shortly after image onset, and latencies did not cover the entire 5 s viewing period. **b**, Neurons did not show much deviation in their response latencies, with the majority having values of a few hundredths of a second. **c**, Most neurons decayed 63% of the way back to baseline firing (Relaxation Time) quickly, but some showed longer relaxation times that spanned the entire 5 s. 23

2.6 **The population of entorhinal neurons encodes time with decreasing accuracy as the interval elapses.** **a**, Performance of a linear discriminant analysis decoder (LDA) trained on the entire population of EC neurons. The x-axis indicates the actual time bin and the y-axis indicates the decoded time, where lighter shading corresponds to higher probabilities. The most frequently decoded time for each actual time bin is marked with a red dot. **b**, Decoding error increases with the passage of time. The x-axis gives the actual time since image onset; the y-axis gives the mean of the absolute error produced by the decoder in **a**. The dotted black line is a fitted regression line. **c**, The temporal code was distributed along many temporal context cells, including those with slow relaxation times. We decoded time from progressively smaller numbers of temporal context cells, by increasing the smallest relaxation time considered. (Bottom left), the absolute error over all time bins as a function of Relaxation Time Threshold. The entire posterior distribution for selected points labeled by Roman numerals are shown along the top of the figure. The threshold is shown as dashed red lines. The horizontal line shows the absolute error in decoding from a permuted dataset; error bars show the 0.025 and 0.975 quantiles. Right: the number of temporal context cells remaining in the analysis as a function of Relaxation Time Threshold. 24

2.7 Time could be decoded above chance from entorhinal firing.

a, A histogram of mean absolute value of decoding error for the permuted data compared to actual decoding error. Training data were the same as the actual decoder, except the true time labels were randomly shuffled. The mean absolute value of decoding error for the original data (.93 s) is marked by a vertical line. **b** A plot of decoder performance compared to permuted data as a function of relaxation time threshold. With 1000 permutations, the significance level is the number of permutations that matched or out performed the actual data divided by 1000 (the line marked by “**” is significant at $p < 0.01$, the line marked by “*” is significant at $p < 0.05$). Thus the decoder is significantly better than chance at $p < 0.01$ up until 2.25 seconds. The decoder is last significantly better than chance at $p < 0.01$ at 3 seconds. This suggests that temporal information is present out to at least 3 seconds. **c** Posterior distributions as bins are dropped. The leftmost sub-panel corresponds to no bins dropped, the next sub-panel corresponds to the first 1.75 seconds dropped the next sub-panel correspond to the first 2.5 seconds dropped, and the rightmost sub-panel corresponds to the first 3.75 seconds dropped) Colorbar and scaling are as in Figure 3a (maximum white correspond to probability 0.5 and greater, minimum black correspond to probability 0.02 and less, color is on log scale).

2.8 **The entorhinal ensemble carries information about stimulus identity.** **a**, We measured the firing rate of each unit to first and second presentations of each image. For each neuron that was recorded for an entire block of images ($n = 270$), we measured the correlation in firing rate across images using Kendall's τ . **b**, The distribution of Kendall's τ for all entorhinal neurons is shown. This distribution is reliably different from zero ($p < 0.001$, see text for details) indicating that entorhinal firing was sensitive to image identity. **c**, Schematic of a population similarity analysis. We measured the similarity of population response to the second presentation of each image to the first presentation of the same image (lag 0). As controls we also computed the similarity between the population response to the second presentation of an image and the responses to the images neighboring the first presentation of that image. Lag -1 refers to the similarity to the immediate predecessor of the image; lag $+1$ refers to the similarity to the immediate successor of the image. **d**, Cosine similarity of the population response to the second presentation of an image to the original presentation of an image (lag 0) and the images neighboring the original presentation of the image (lags -1 and $+1$) over 64 blocks of first and second presentations. The similarity to lag 0 is greater than the similarity to either of the neighboring images. *** indicates significance at the $p < 0.001$ level. Error bars correspond to the 95 % confidence interval of mean cosine similarity calculated across sessions. 31

3.1 Distinguishing two potential accounts for the recency effect from response time distributions. **Top.**

In strength-based accounts of recency, all items are available simultaneously but older items (darker lines) are represented with less fidelity than newer items (lighter lines). *Left:* Probability density functions of response times for each lag. The shape of the distribution changes with the drift rate. *Middle:* A cumulative distribution of response times for different lags. RT distributions start at the same point regardless of lag, but spread as you move into later quantiles. *Right:* A plot of lag modulation factor as a function of quantile. In strength-based accounts of recency, lag modulation factor is zero at the start of the distribution and monotonically increases for later quantiles. **Bottom.** If the recency effect arises due to a delay in recovering older temporal contexts, then older items require more time before evidence accumulation can begin. *Left:* Probability density functions of response times for each lag. Rather than varying in their drift rate, newer items begin accumulating evidence before older items. *Middle:* A cumulative distribution of response times for different lags. Because the distributions do not change their shape, the distance between response times is similar across deciles. *Right:* The Lag Modulation Factor as a function of quantile. The change in the start to accumulate evidence results in a non-zero intercept. . . .

3.2	The recency effect is robust across experimental conditions.	
	<p>a. Hit Rate as a function of \log_2 lag for Experiments 1-6. The hit rate goes down with lag. There is a more pronounced drop in the hit rate at higher lags. b. Median response time as a function of \log_2 lag. Median response time increased linearly with the logarithm of lag. c. The 1st decile as a function of \log_2 lag. 1st decile response times increased linearly with the logarithm of the lag at a similar rate as median response times. Error bars in all figures represent the 95% confidence interval of the mean across participants normalized using the method described in Morey (2008).</p>	63
3.3	Time to access memory changed systematically with lag in all six experiments.	
	<p>Unsmoothed across-participant cumulative response distributions for each lag. Shorter lags correspond to darker lines. Note that the cumulative distributions shift with decreasing recency. The inset consists of a zoomed in view of the rising points of the distributions. For all six experiments, there is consistent evidence that newer items are available before older items.</p>	65
3.4	Recency impacted response times for even the fastest responses.	
	<p>Histograms of each subject's slope per lag (ms) for the fastest responses as calculated by the Lag Modulation Factor analysis. Across all six experiments, subjects had significant recency effects at the start of their RT distributions.</p>	68

3.5	The effect of recency on response times was the same regardless of repetitions. Median response time as a function of the most recent lag and number of repetitions. To the extent the lines are parallel, it means that the effect of recency on median RT's was the same for across repetitions. Analyses reported in the text demonstrated that only the most recent lag affected the RT of old probes repeated twice.	74
4.1	A boxplot of median number of words recalled per trial across participants and presentation rates, with interquartile range, 95% confidence intervals and outliers. Participants recalled fewer words as the presentation rate increased.	82
4.2	Probability of first recall. Participants tended to begin recall by naming an item from the beginning or end of the list. As presentation rates increased, the probability of initiating recall at the end of the list decreased, and the probability of initiating recall with at the beginning of the list increased.	83
4.3	Probability of recall as a function of position in study list. Participants showed the highest level of performance for items at the beginning and end of the list. As presentation rate increased, participants showed a tendency to have a lower recency effect in comparison to the primacy effect.	84

4.4	Matrices showing the number of total valid recall transitions between any two study list positions for each presentation rate separately. Colors and numbers correspond to the number of such recalls summed across participants. Transitions between extreme positions in study lists correspond to primacy and recency effects, which persist across rates. In contrast, the likelihood of recalling nearby items (i.e., close to the diagonal) appears to decrease as presentation rate increases.	84
4.5	Lag-CRP for transitions restricted to serial positions 7-13 to avoid confounds with primacy and recency effects. As presentation rates increase, the contiguity effect weakens but remains asymmetric. At 8 Hz there is no positive evidence for a contiguity effect.	86

List of Abbreviations

CRP	Conditional Response Probability
EC	Entorhinal Cortex
hz	Hertz
iEEG	Intracranial Electroencephalography
LEC	Lateral Entorhinal Cortex
LDA	Linear Discriminant Analysis
MEG	Magnetoencephalography
ms	Milliseconds
MTL	Medial Temporal Lobe
PBE	Population Burst Event
PFR	Probability of First Recall
PSTH	Peristimulus Trial Histogram
RSVP	Rapid Serial Visual Presentation
RT	Response Time
SPC	Serial Position Curve
SPW-R	Sharp-wave Ripple

Chapter 1

Introduction

Episodic memory, memory for an event as well as where and when it occurred - is believed to depend on the recovery of its associated context (Tulving, 1983). To better understand the idea of episodic memory and retrieving an associated context, picture yourself walking over “alphabet hill” as illustrated in Figure 1.1. As you walk, you take note of each letter. The thought bubble represents your recent past or context. As you move along the hill, your recent past changes. Later, if asked “did you see the letter C on your walk”, you can think back and remember that you did. The letter C is not the only thing that comes to mind in this scenario. You recover the entire context associated with C. Formalizing this idea of recovering spatiotemporal context, mathematical psychologists proposed retrieved-context theories (Howard & Kahana, 2002; Polyn, Norman, & Kahana, 2009; Lohnas, Polyn, & Kahana, 2015) to make sense of a wide range of behavioral findings, particularly the recency and contiguity effects.

1.1 Behavioral effects in memory

A hallmark of memory is that, all else equal, more recent events are more likely to be remembered. This phenomenon, the recency effect, is apparent in many memory tasks. This is apparent in immediate free recall, a paradigm where individuals study a list of items and begin recall as soon as the list ends. In immediate free recall, the final few items in the list are more likely to be recalled (B. B. Murdock,

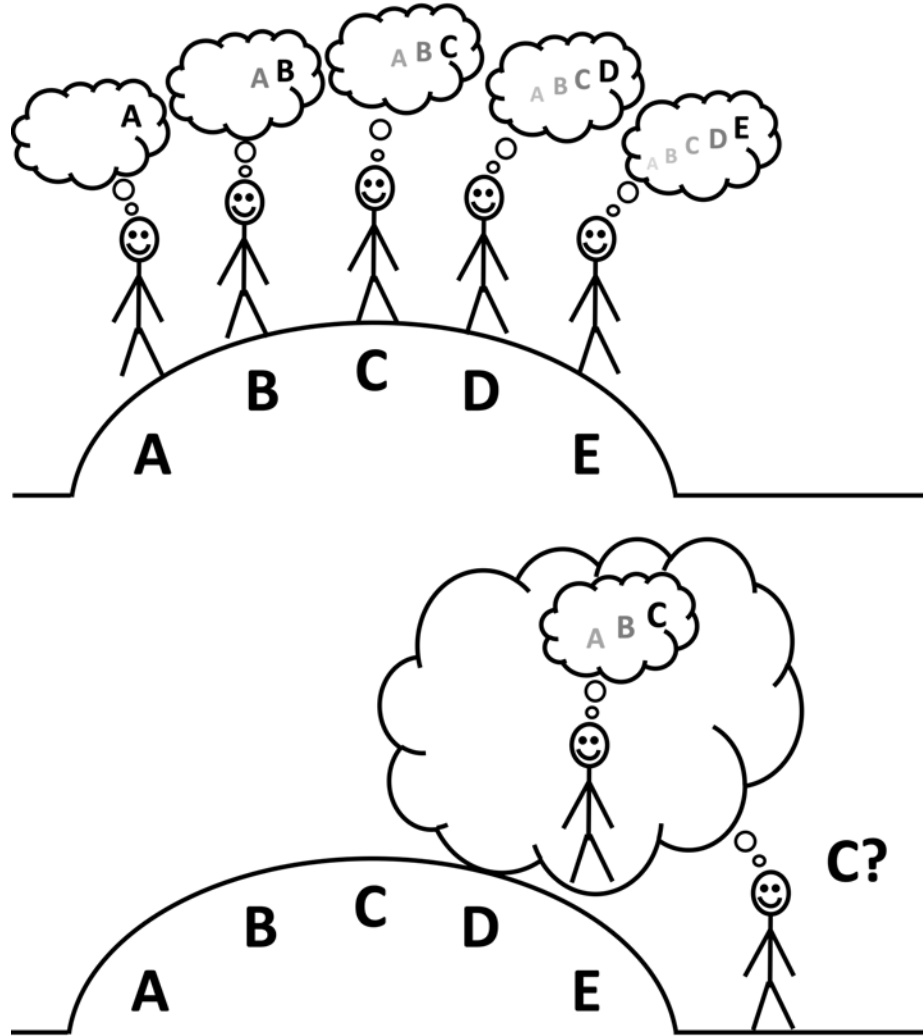


Figure 1.1: A walk along alphabet hill. *Top:* An illustration of an individual walking along a hill with letters on it. At each moment, the thoughts of the individual, represented as thought bubbles, includes the recent past. This is their temporal context (e.g., A, B, C). *Bottom:* Jumping back in time. When asked to remember if they saw the letter ‘C’, the individual retrieves the entire context associated with ‘C’. Adapted from Folkerts et al., 2018.

1962; Glanzer & Cunitz, 1966). In recognition tasks, the recency effect manifests as higher hit rates and faster response times for recently presented items (Shepard & Teghtsoonian, 1961; B. B. Murdock & Anderson, 1975; Monsell, 1978; Hockley, 1982; Donkin & Nosofsky, 2012). In paired-associate tasks, the likelihood of recalling a given pair decreases as the amount of time/number of items presented between study and retrieval increases (Rubin, Hinton, & Wenzel, 1999). The recency effect is also present in more naturalistic tasks, such as remembering where one parked their car (da Costa Pinto & Baddeley, 1991) or recalling autobiographical events (Moreton & Ward, 2010).

Another common phenomenon in episodic memory is the contiguity effect, the finding that after recollecting something, other items experienced “close” in time are more likely to be recalled next (Kahana, 1996; G. Schwartz, Howard, Jing, & Kahana, 2005; Davis, Rizzuto, Geller, & Kahana, 2008). The contiguity effect manifests in how recall unfolds over time in free recall. After recalling a word from the list, the next word recalled is more likely to come from adjacent places on the list (Kahana, 1996; Howard & Kahana, 1999; Howard, Youker, & Venkatadass, 2008; Healey, Long, & Kahana, 2019). In addition to the tendency to recall nearby items, forward transitions are more likely than backward transitions of the same distance on the list. This produces a pronounced asymmetry in conditional-response probability. Analogous decreases in inter-response times also occur in free recall (Kahana, 1996; Healey et al., 2019). However, the contiguity effect is not limited to lists of words, as it occurs when recalling news events (Uitvlugt & Healey, 2019), or autobiographical events (Moreton & Ward, 2010). In recognition tasks, for a repeated item that follows a correct high-confidence old response, there is an increase in hit rate and decrease in response time if the repeated item was studied nearby in time to the just recognized item (G. Schwartz et al., 2005; Averell, Prince, & Heathcote, 2016; Zhou, Osth, &

Smith, 2023). The contiguity effect in recognition does not appear to be asymmetric. Contiguity effects are also seen in the patterns of errors made in paired associate (Davis et al., 2008) and serial recall tasks (Kahana & Caplan, 2002; Solway, Murdock, & Kahana, 2012).

1.2 Retrieved context models

In light of these recency and contiguity effects, memory researchers proposed many classes of models to account for these effects, including two-store models (Atkinson & Shiffrin, 1968; Raaijmakers & Shiffrin, 1980) and distributed memory models (B. B. Murdock, 1982; Gillund & Shiffrin, 1984; D. Hintzman, 1987; Shiffrin & Steyvers, 1997). One class of models that have succeeded in accounting for recency and contiguity are retrieved-context models (Howard & Kahana, 2002; Sederberg, Howard, & Kahana, 2008; Polyn et al., 2009; Lohnas et al., 2015). Building on earlier variable context models (Estes, 1955; J. R. Anderson & Bower, 1972; Mensink & Raaijmakers, 1989), these models argue that context consists of a distributed representation that drifts over time. As events occur, they form associations with the present state of context. These associations are formed in both directions; context can serve as a cue for an event, and events can serve as a cue for its context. When recalling an event or experiencing it again, its associated context is reinstated. Context can also drive the recall of associated items. As a result, the likelihood of remembering an event depends on the similarity between the present context and the context when the event first occurred. The more similar the prior context is to the current context, the more likely recovery is. Context serves as the primary cue driving the recall of events. Retrieved-context theories account can account for the contiguity and recency effects described above via their slowly drifting context.

In these models, the context for items experienced recently should be similar to the

present context. This high overlap in context for recently experienced items increases their probability of being remembered. In Figure 1·1, the thought bubble when at ‘E’ is more similar to the bubble at ‘D’ than ‘B.’ These theories can also account for contiguity effects through their retrieval mechanisms. In retrieved-context models, recalling a memory reinstates its associated context. This reinstated context should overlap with the context for other items that occurred “close” in time, boosting their likelihood of being recalled next. For example, upon recalling ‘C’ in Figure 1·1, your reinstated context looks more like “B” than “A”.

Empirical findings in the medial temporal lobe (MTL), a region critical to episodic memory function (Scoville & Milner, 1957; Eichenbaum, Yonelinas, & Ranganath, 2007; Dede, Frascino, Wixted, & Squire, 2016), offer support to the retrieved-context account of recency and contiguity. As expected from a slowly drifting context representation that would produce a recency effect, firing patterns within MTL structures drift over time such that the similarity between the pattern of activity at any two points in time decreases as their distance in time increases (Manns, Howard, & Eichenbaum, 2007; Mankin et al., 2012; Howard, Viskontas, Shankar, & Fried, 2012; Hsieh, Gruber, Jenkins, & Ranganath, 2014; Folkerts, Rutishauser, & Howard, 2018; Mau et al., 2018; Liu et al., 2022). Because multiple experimental techniques (i.e., single-unit recordings, calcium imaging, and fMRI) have found drift in the MTL, it is unlikely that drift is simply a recording artifact. There is also increasing experimental support that the MTL jumps back in time and recovers the prior context during episodic memory retrieval. Following episodic memory retrieval, the MTL reinstates the pattern of activity present at encoding (Manning, Polyn, Litt, Baltuch, & Kahana, 2011; Howard et al., 2012; Yaffe et al., 2014; Folkerts et al., 2018). Because the activity drifts slowly over time, this reinstated context is similar to contexts for nearby events, resulting in a neural contiguity effect.

1.3 A scale-invariant representation of temporal context

Although the retrieved-context models described above have successfully described behavioral results and have found empirical support in the brain, their formulation has fundamental issues. First, recency and contiguity manifest in behavior over a wide range of time scales (Bjork & Whitten, 1974; Glenberg et al., 1980; Howard & Kahana, 1999; Howard et al., 2008; Unsworth, 2008; Moreton & Ward, 2010; Uitvlugt & Healey, 2019), suggesting that the memory representation must be scale-invariant (Chater & Brown, 2008). The system’s output should be consistent with a rescaled input. In standard retrieved-context models (Howard & Kahana, 1999; Polyn et al., 2009; Logan, 2018), a single exponential decay drives context drift. This exponential decay forces a time scale on the drift. If events happen at rates faster than the decay rate, context does not drift much between items. If the spacing between events is longer than the decay rate, then the context for the previous item fully decays before the next item appears. In both cases, a single exponential decay would not be able to capture behavioral effects accurately (Howard, 2022).

Given these results, a scale-invariant form of context was proposed, which can produce recency and contiguity effects at multiple scales (Shankar & Howard, 2012; Howard et al., 2014; Liu, Tiganj, Hasselmo, & Howard, 2019; Howard & Hasselmo, 2020). In this framework, a two-layer feed-forward network represents temporal context. The first layer, F , contains a population of leaky integrators that fire immediately after an event and then exponentially decay back to baseline. Critically, these leaky integrators have a log-spaced spectrum of decay rates, s , rather than a single decay rate as in standard retrieved-context models. The output of this first layer F feeds into a second layer of cells that fire sequentially, labeled T . In this population, rather than a log-spaced spectrum of decay rates s , the peak times τ^* are log-spaced, so more cells peak early in the delay later. This distribution of peaks forms a set of

temporal receptive fields that are optimally distributed for situations in which there is no characteristic time scale (Shankar & Howard, 2013), as is believed to be the case for episodic memory (Glenberg et al., 1980; Howard et al., 2008). This class of models can account for contiguity and recency effects in behavior at multiple time scales (Howard, Shankar, Aue, & Criss, 2015). In this framework, contextual overlap falls off according to a power law, which does not have an inherent time scale. Overlap between events decreases logarithmically, such that the difference between 4 s and 6 s ago is more like the difference between 40 s and 60 s than between 40 s and 42 s. This smooth change enables the model to account for recency and contiguity at all time scales simultaneously. The degree to which items are “nearby” on the scale of seconds to minutes to hours becomes similar.

1.4 Neural representations of time

This scale-invariant representation of context makes several empirical predictions about the brain and behavior. In particular, it predicts two forms of temporal representations that should exist in the brain. First, the sequentially activated cells in the second layer of this model align well with time cells. Time cells, cells with receptive fields in time following some event, were first reported in rodent hippocampus (Pastalkova, Itskov, Amarasingham, & Buzsaki, 2008; MacDonald, Lepage, Eden, & Eichenbaum, 2011; MacDonald, Carrow, Place, & Eichenbaum, 2013). Subsequent work has found time cells in additional regions of the rodent brain, such as medial entorhinal cortex (Kraus, Robinson, White, Eichenbaum, & Hasselmo, 2013), striatum (Mello, Soares, & Paton, 2015; Akhlaghpour et al., 2016), and prefrontal cortex (Tiganj, Kim, Jung, & Howard, 2017). Time cells are not exclusive to the rodent brain; they have been reported in monkey hippocampus (Cruzado, Tiganj, Brincat, Miller, & Howard, 2020) and prefrontal cortex (Tiganj, Cromer, Roy, Miller,

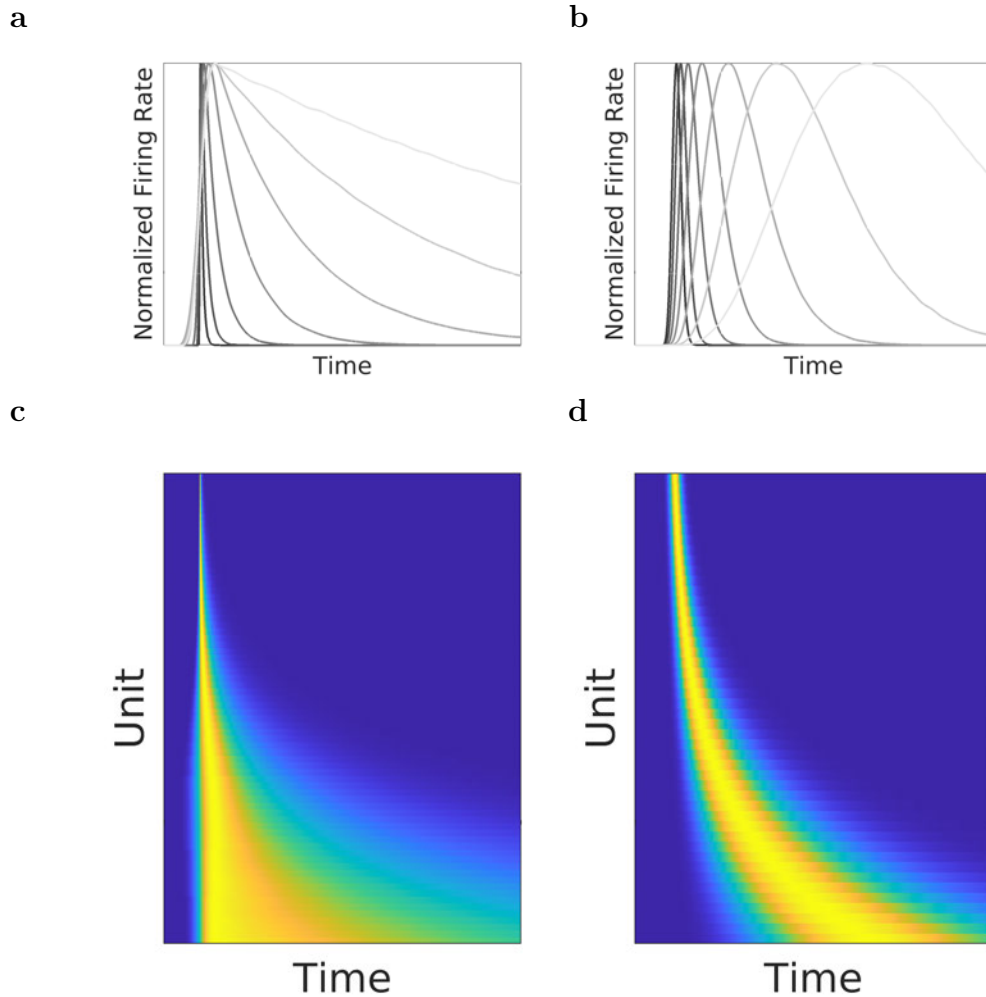


Figure 1.2: Scale-invariant representations of temporal context. **a.** Peristimulus time histograms (PSTH) of hypothetical cells in the first layer F . Lighter lines correspond to cells with longer relaxation times s . **b.** PSTH of hypothetical cells in the second layer T . Lighter lines correspond to cells with later peak times τ^* . **c.** Heatplot of a population of hypothetical cells in F , sorted by their corresponding relaxation time, where yellow corresponds to high firing and blue low firing. **d.** Same format as **c** but for hypothetical cells in T sorted by their corresponding peak time.

& Howard, 2017; Cruzado et al., 2020) as well as in the human medial temporal lobe (Reddy et al., 2021; Schonhaut, Aghajian, Kahana, & Fried, 2022). As would be expected from scale-invariant sequential activity, the distribution of peak times in a population of rodent hippocampal time cells is distributed according to a power law distribution (Cao, Bladon, Charczynski, Hasselmo, & Howard, 2021). Further, time cells that fire later in a delay have wider time fields (Salz et al., 2016; Tiganj, Cromer, Roy, Miller, & Howard, 2018; Cruzado et al., 2020; Cao et al., 2021).

The cells in the first layer of the scale-invariant model should respond immediately after an event and then monotonically relax their firing at a spectrum of rates. Indeed, there is evidence of cells that monotonically change their firing in various brain regions (Leon & Shadlen, 2003; Lebedev, O’Doherty, & Nicolelis, 2008; Kim, Ghim, Lee, & Jung, 2013; Guo, Huson, Macosko, & Regehr, 2021), including the medial temporal lobe (Naya & Suzuki, 2011; Tsao et al., 2018; Ning, Bladon, & Hasselmo, 2022). What has been underexplored is if these monotonic cells have a spectrum of time constants controlling their change in firing (but see Tsao et al., 2018; Guo et al., 2021). In order to form a scale-invariant representation of context, the distribution of time constants should come from a power law distribution analogous to the distribution of peak times in time cells (Cao et al., 2021). I refer to cells in this first layer as “temporal context cells”.

1.5 Overview of chapters

Compared to time cells, there have been few reports of the properties and existence of temporal context cells. Given that the entorhinal cortex projects directly to the hippocampus (Witter, Wouterlood, Naber, & Van Haeften, 2000), this region has a high potential to contain temporal context cells. Indeed the most substantial evidence for temporal context cells comes from rodent lateral entorhinal cortex. Tsao et al. (2018)

reported cells that responded shortly after beginning a session or trial and then relaxed their firing with a spectrum of rates spanning minutes to hours in rodent lateral entorhinal cortex. Because the monkey hippocampus contains time cells (Cruzado et al., 2020), one would expect to find temporal context cells in the monkey entorhinal cortex. Further, although consistent with a scale-invariant representation, the time scales reported in rodent lateral entorhinal cortex are far longer than time cells, which fire on the scale of seconds. The model predicts that there should also be temporal context cells that relax their firing on the scale of seconds, as is the case with time cells. Chapter 2 presents evidence of temporal context cells in the monkey entorhinal cortex. While freely viewing images, this population of units responds shortly after the image appears on screen and then relaxes their firing over seconds with different rates.

If temporal context is indeed scale-invariant, then the distance between any two contexts is a logarithmic function of their separation in time. This scaling implies that the time to reinstate a prior temporal context should increase logarithmically. In continuous recognition, previous work has suggested that the recency effect is logarithmic out to at least 4 minutes (Hockley, 1982). Increasing the time to reinstate older contexts can account for this effect. Chapter 3 presents a series of 6 continuous recognition experiments. Response times increased in all experiments logarithmically out to at least 8 minutes. Careful analysis of the distribution of response times indicated that this increase arises from a delay in reinstating older contexts.

Finally, the fastest exponential decay rate is a free parameter in this scale-invariant model. The behavioral contiguity is incredibly robust, regardless of most methodological manipulations (Healey et al., 2019). Finding a presentation rate at which contiguity breaks down would suggest a lower limit on the decay rates. The rapid serial visual presentation literature suggests a boundary condition on the contiguity

effect. For lists of items presented at 10 Hz, items are no longer processed separately but are processed together as a single event (Potter & Levy, 1969). As a result, at fast presentation rates, it is unlikely that each item would be bound to a temporal context, disrupting the behavioral contiguity effect. Chapter 4 presents behavioral data in humans indicating that presenting items at 8 Hz but not 2 or 4 Hz disrupts the contiguity effect in free recall.

Chapter 2

A temporal record of the past with a spectrum of time constants in the monkey entorhinal cortex.

This chapter is based on previously published work, reported in Bright et al., 2020.

2.1 Introduction

Episodic memory, the vivid recollection of an event situated in a specific time and place (Tulving, 1983), depends critically on medial temporal lobe (MTL) structures, including the hippocampus and entorhinal cortex (EC) (Milner, 1959; Eichenbaum et al., 2007; Dede et al., 2016; Squire, Stark, & Clark, 2004). Building on pioneering work demonstrating a spatial code in the hippocampus and entorhinal cortex (O'Keefe & Dostrovsky, 1971; Fyhn, Molden, Witter, Moser, & Moser, 2004), recent research has shown that hippocampal representations also carry information about the time at which past events took place, suggesting that the MTL maintains a representation of spatiotemporal context in support of episodic memory (Pastalkova et al., 2008; MacDonald et al., 2011; Eichenbaum, 2017). Although a great deal is known about the temporal coding properties of neurons in the hippocampus, the temporal code in the entorhinal cortex, which provides the majority of the cortical input to the hippocampus is less understood, but see (Naya & Suzuki, 2011; Kraus et al., 2015; Naya, Chen, Yang, & Suzuki, 2017; Tsao et al., 2018).

Hippocampal time cells provide a record of recent events including explicit information about when an event occurred. Analogous to hippocampal place cells that fire when an animal is in a circumscribed region of physical space (Wilson & McNaughton, 1993; O’Keefe & Dostrovsky, 1971), hippocampal time cells fire during a circumscribed period of time within unfilled delays (Pastalkova et al., 2008; MacDonald et al., 2011; Kraus et al., 2013). Across studies, there is a remarkable consistency in the properties of hippocampal time cells. Hippocampal time cells peak at a range of times during the delay interval and typically code time with decreasing accuracy as the delay unfolds, as manifest by fewer neurons with peak firing late in the delay and wider time fields later in the delay (Kraus et al., 2015; Salz et al., 2016; Mau et al., 2018). Hippocampal time cells have been observed in a wide range of behavioral paradigms, including tasks with and without explicit memory demands during the delay (Salz et al., 2016) and experiments in which the animal is fixed in space (MacDonald et al., 2013; Terada, Sakurai, Nakahara, & Fujisawa, 2017). In addition, it has been shown that different stimuli trigger different time cell sequences (Terada et al., 2017; MacDonald et al., 2013). Taken together, time cells provide an explicit record of how far in the past an event took place, i.e., the amount of time that has passed since the beginning of a delay period or since the presentation of a to-be-remembered stimulus. By examining which time cells are active at a particular time, we can easily determine not only what event took place, but how far in the past that event occurred.

Many of the properties of hippocampal time cells have been observed in other brain regions including prefrontal cortex (Bolkan et al., 2017; Tiganj, Kim, et al., 2017; Tiganj et al., 2018; Jin, Fujii, & Graybiel, 2009) and striatum (Jin et al., 2009; Mello et al., 2015; Akhlaghpour et al., 2016) suggesting that the hippocampus is part of a widespread network that carries episodic information. A recent report from the

rat lateral EC adds important data to this growing body of literature regarding the representation of time in the brain. Tsao et al. observed a population of neurons that changed slowly and reliably enough to decode time within the experiment over a range of time scales (Tsao et al., 2018). Unlike time cells, which respond a characteristic time since the event that triggers their firing, lateral EC neurons respond immediately upon entry into a new environment, and then relax slowly. The relaxation times across individual neurons were very different, ranging from tens of seconds to thousands of seconds. To distinguish this population from time cells we will refer to neurons that are activated by an event and then relax their firing gradually as *temporal context cells*. The designation “temporal context cells” is not meant to indicate some intrinsic biological property but is simply intended, much like the terms “place cell” or “time cell” as a convenient short-hand to describe the functional properties of these neurons. Because these temporal context cells code for time, but with very different properties than time cells, these two populations provide a potentially important clue about the nature of temporal coding in the brain and the neural mechanisms that may support episodic memory.

Here, we identified temporal context cells in monkey EC during a free-viewing task (Meister & Buffalo, 2018). We examined EC neuron responses in a five-second period after presentation of an image. In the time after presentation of the image, a representation of what happened when should carry both time-varying information about when the image was presented, as well as information that discriminates the identity of the image. To anticipate the results, the data demonstrate that neurons in monkey EC are activated shortly after a visual stimulus and then decay with a variety of rates, enabling reconstruction of when the image was presented. This form of temporal coding is similar to temporal context cells observed in rat lateral EC (Tsao et al., 2018), but is qualitatively different from time cells that have been observed in

the hippocampus and other regions. Because each image was shown twice over the course of the experiment, we were able to assess whether the pattern of activation over neurons depends on the identity of the image presented. Taken together, these data suggest that these temporal context cells carry information about what happened in addition to when it happened.

2.2 Results

A total of 349 neurons were recorded from the entorhinal cortex (EC) in two macaque monkeys during performance of a visual free-viewing task. Each trial began with a required fixation on a small cross, followed by the presentation of a large, complex image that remained on the screen for five seconds of free viewing (Figure 2·1a). Unlike canonical hippocampal time cells, which are activated at a variety of points within a time interval (e.g., Figure 2·1c), most entorhinal neurons changed their firing a short time after the presentation of the image. Figure 2·3a shows three representative neurons that responded to image presentation (more examples are shown in Figure 2·4). While most of these neurons increased their firing rate after the image was presented, some neurons decreased their firing rate in response to image presentation. Although behavior was not controlled during the five second free-viewing period, the response of these neurons was consistent across trials, which can be seen by examination of the trial rasters, indicating that the temporal responsiveness was unlikely to be a correlate of behavior.

Although the image-responsive neurons in EC responded at about the same time post-stimulus, they relaxed back to their baseline firing at different rates. Whereas some neurons relaxed back to baseline quickly (Figure 2·3a, left), some relaxed much more slowly. For instance the neuron shown in the right of Figure 2·3a did not return to baseline even after five seconds.

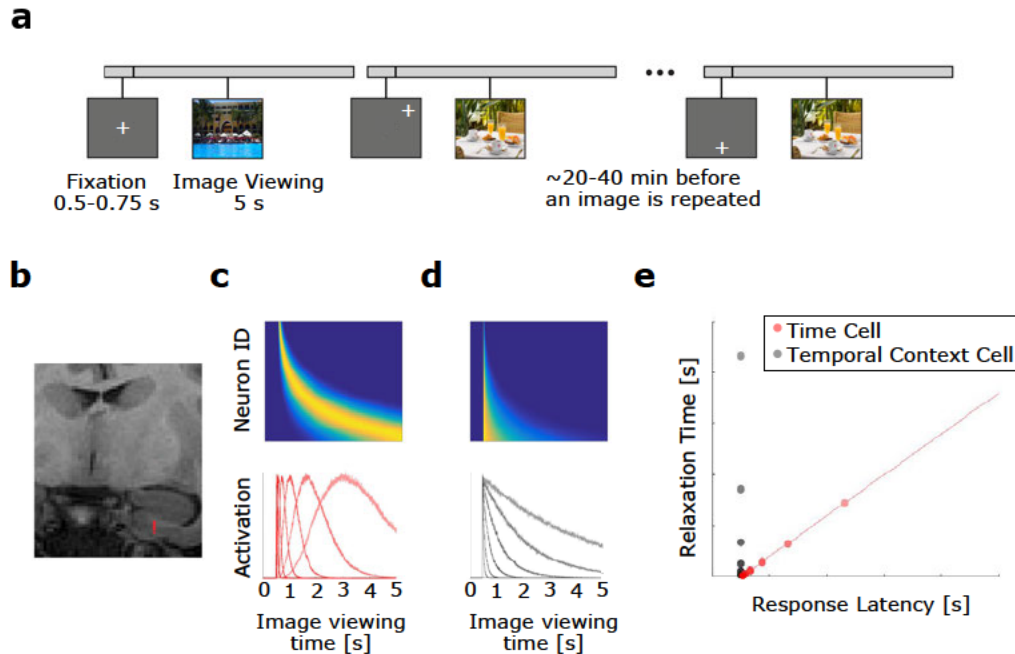


Figure 2.1: a-b: Experimental procedures. a, On each trial, the monkey freely viewed an image. After the monkey viewed an image for 5 seconds, the image disappeared. Following every trial, the monkey performed gaze calibration trials and received a reward. Images were presented twice during an experimental session. b, Estimated recording channel location in the entorhinal cortex for one recording session is shown in red on a coronal MRI. **c-e: Two hypotheses for neural representations of time.** c-d, Heat plot (top) and tuning curves (bottom) for two hypotheses for how a time interval of image viewing could be represented. In the heat plots, cooler colors correspond to low activity while warmer colors correspond to higher activity. c, Hypothetical activity for sequentially activated time cells, like those observed in the hippocampus. Because the time of peak response across neurons covaries with the spread of the firing field, neurons with later firing fields display wider firing fields. d, Hypothetical activity for monotonically decaying temporal context cells, like those observed in rodent EC. Neurons in this population reach their peak at about the same time. However, different neurons decay at different rates. e, Properties of neurons representing time passage by the hypothetical cells in c and d. A population of time cells (red) should respond at different points in a time interval, and show a robust correlation between response latency and reaction time. Conversely, a population of temporal context cells (gray) should respond in a more restricted time range shortly after the start of a time interval, and show no correlation between response latencies and relaxation times. baseline.

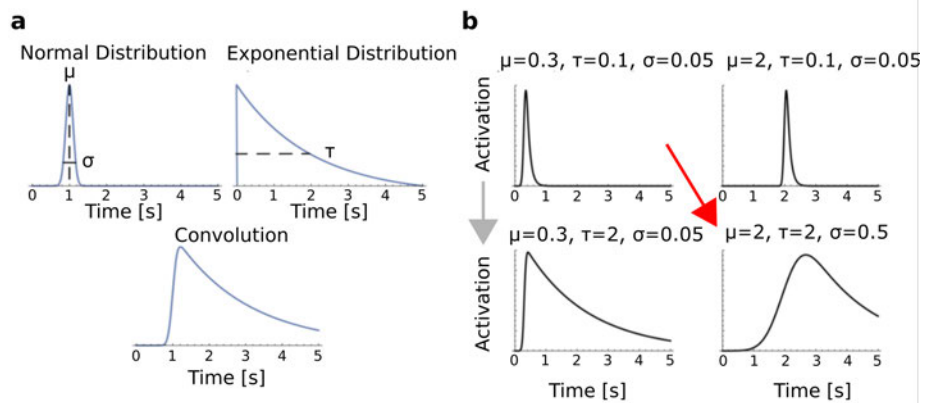


Figure 2-2: Modeling neuron responses using an ex-Gaussian distribution. **a**, The ex-Gaussian model of a neuron's response (bottom) is formed by convolving a Gaussian distribution (with parameters μ and σ) with an exponential distribution (with parameter τ) (top). **b**, In the ex-Gaussian model, an increase in μ shifts the distribution to the right, an increase in σ widens the central peak of the distribution, and an increase in τ lengthens its decay rate. The gray and red arrows correspond to the parameter changes expected in the process of fitting the response of temporal context cells and sequentially activated time cells respectively.

2.2.1 Temporal receptive field analysis revealed a population of temporal context cells in entorhinal cortex

To quantify the qualitative results from examination of the individual units, we classified neurons that changed their firing in synchrony with image presentation and measured their temporal receptive fields with a model with three parameters. Each neuron’s firing rate in the time from 0.5 s before image onset to 5 s after was quantified using a model-based approach. The firing field was estimated as a convolution of a Gaussian (latency in neuronal response) and an exponential decay (Figure 2.2a). This approach builds on previous work to estimate time cell activity as a Gaussian firing field (Tiganj, Kim, et al., 2017; Salz et al., 2016; Tiganj et al., 2018). The method for estimating parameters is described in detail in the methods. In this model, we were able to quantify apparent response properties using two key parameters: 1) the parameter μ , which describes the mean of the Gaussian, estimates the time at which each neuron begins to respond (Response Latency) and 2) the parameter τ , which describes the time constant of the exponential term, estimates how long each neuron takes to relax back to 63% of its maximum deviation from baseline firing (Relaxation Time) (Figure 2.2b). A third parameter σ controls the standard deviation of the Gaussian term. Using this model, we tested whether a neuron had a time-locked response to image onset by quantifying the extent a model with a temporal response field fit the neuron’s data better than a model with only constant firing including the prestimulus period. Neurons that were better fit by including a temporal firing field were referred to as “visually responsive.” Importantly, this method would identify populations of either hippocampal time cells and temporal context cells as visually responsive. However, as shown in Figure 2.1e, these two different forms of temporal coding would produce distinguishable distributions of parameters.

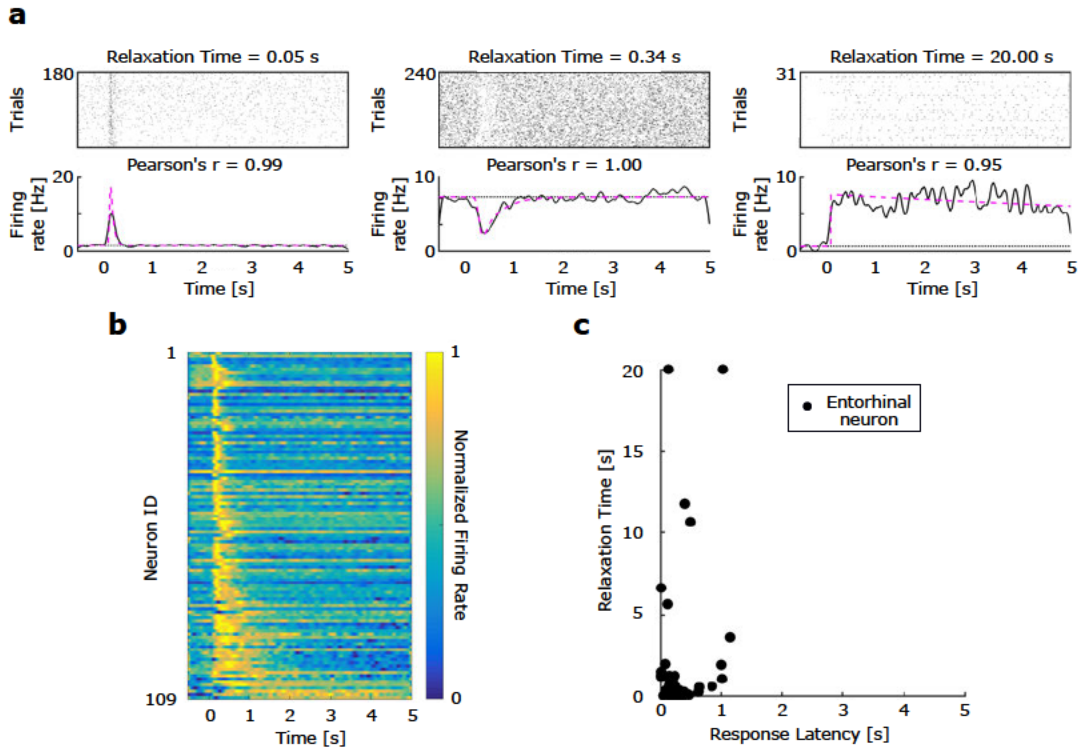


Figure 2-3: Temporal context cells in monkey entorhinal cortex respond shortly after image presentation, then return to baseline at a spectrum of rates. **a**, Three representative cells that responded to image onset and decayed at different rates (Figure 2-4 shows more examples). Each pair of plots indicate the activity of one neuron relative to image onset in a raster plot (top) and PSTH (bottom). In each raster, a tick mark indicates when the neuron fired an action potential. In the PSTH, the solid black line indicates the smoothed firing rate, the pink line indicates the model estimate of firing rate, and the estimated baseline firing rate is indicated by the black dotted line. Relaxation Time refers to the duration between response peak and when the neurons returned 63% of the way to baseline. Pearson's r is the correlation between the fits for even and odd trials. **b**, A heatplot of the normalized firing rate of 109 temporal context cells, relative to image onset, sorted by their Relaxation Time. The color scheme is the same as in Figure 2-1c, d. The majority of neurons responded within 1 s of image onset, relaxing back to baseline with a spectrum of decay rates. **c**, A scatter plot of the joint distribution of each neuron's Response Latency and Relaxation Time. Response Latencies did not span the entire 5 s, unlike Relaxation Time, and a neuron's Response Latency and Relaxation Time were not correlated. Figure 2-5 shows the marginal distributions of these parameters.

A substantial fraction of entorhinal neurons changed their firing in response to image presentation

In order to minimize the noise and obtain the most accurate distribution of response parameters across neurons, we used a conservative criterion to identify neurons that responded to image presentation. This method identified 109/349 neurons as visually responsive. Of those 109 responsive neurons, 84 neurons showed an increase in their firing rate in response to image onset, whereas 25 showed a decrease in their firing rate. Figure 2-3b summarizes the temporal response properties of these 109 neurons. Each row of the figure shows the averaged response of one neuron over the course of a trial. The data demonstrate that almost all of the neurons reached their maximum deviation from baseline within a few hundred milliseconds of the image presentation. This can be appreciated from the vertical yellow strip along the left edge of the heat map. These results are in striking contrast to the typical responses of hippocampal time cells (e.g., Figure 2-1c). Analogous plots for hippocampal time cells, which vary smoothly in their peak times, result in a curved ridge extending from the upper left to the lower right. In contrast, the variability across neurons in this entorhinal population was not in the time point at which the neurons reached their maximum deviation from baseline, but rather in the time course over which each neuron relaxed. This can be appreciated in the progressive widening of the ridge in Figure 2-3b from top to bottom.

Visually responsive entorhinal units showed short Response Times but a broad distribution of Relaxation Times

Figure 2-3c shows the Response Latency and Relaxation Time for the 109 entorhinal neurons that were categorized as visually responsive (Figure 2-5 shows the marginal distributions for each parameter). Response Latency values were clustered tightly at small values (median = 0.16 s, interquartile range = 0.13 s to 0.24 s). For 90% of

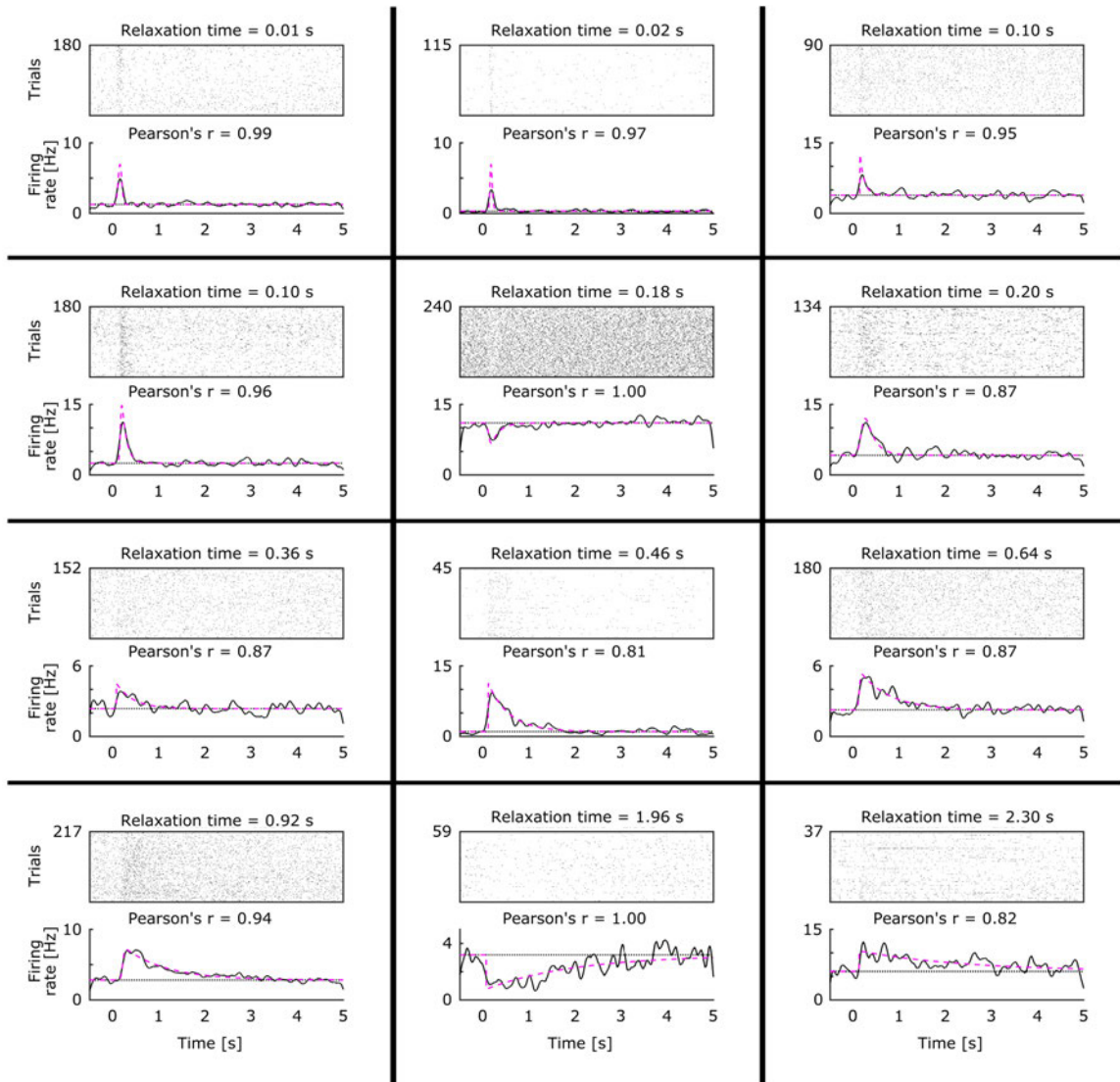


Figure 2.4: Additional examples of temporal context cells. Format is as in Figure 2.3a.

neurons, the Response Latency was less than 0.40 s. In contrast, Relaxation Times showed a wider distribution (median = 0.23 s, interquartile range = 0.10 s to 0.61 s, 90th percentile = 1.29 s), and even included values longer than the 5 s duration of the viewing period. Lastly, the third parameter σ , which controls the standard deviation of the Gaussian, was small and tightly clustered across neurons (median = 0.02 s, interquartile range of 0.001 s to 0.06 s, 90th percentile = 0.31 s). The consistently small value of this parameter indicates that the shape of the temporal receptive fields was well-described by a delayed exponential function.

Response Time and Relaxation Time were not correlated across neurons

Across the 109 neurons, Response Latency and Relaxation Time were not significantly correlated with one another, Kendall's $\tau = 0.03$, $p = 0.64$. To assess whether this null effect was reliable, we computed the Bayes factor, which enables an estimate of the likelihood of the null hypothesis. This analysis yielded a Bayes factor of $\text{BF}_{01} = 7.17$, providing support that neuron Response Latency and Relaxation Time values are uncorrelated. Across neurons, Response Latency and σ were also not correlated with one another, Kendall's $\tau = -0.04$, $p = 0.59$, $\text{BF}_{01} = 6.87$. Unlike hippocampal time cells, there was no evidence that temporal context cells that peaked later in the time interval showed broader firing fields. In contrast to hippocampal time cells, which show a systematic relationship between the peak time of firing and the width of the temporal firing field (Kraus et al., 2013; Howard et al., 2014; Salz et al., 2016), the overarching conclusion from these analyses is that the firing of entorhinal neurons deviated from background firing shortly after the presentation of the stimulus and then relaxed exponentially at a variety of rates.

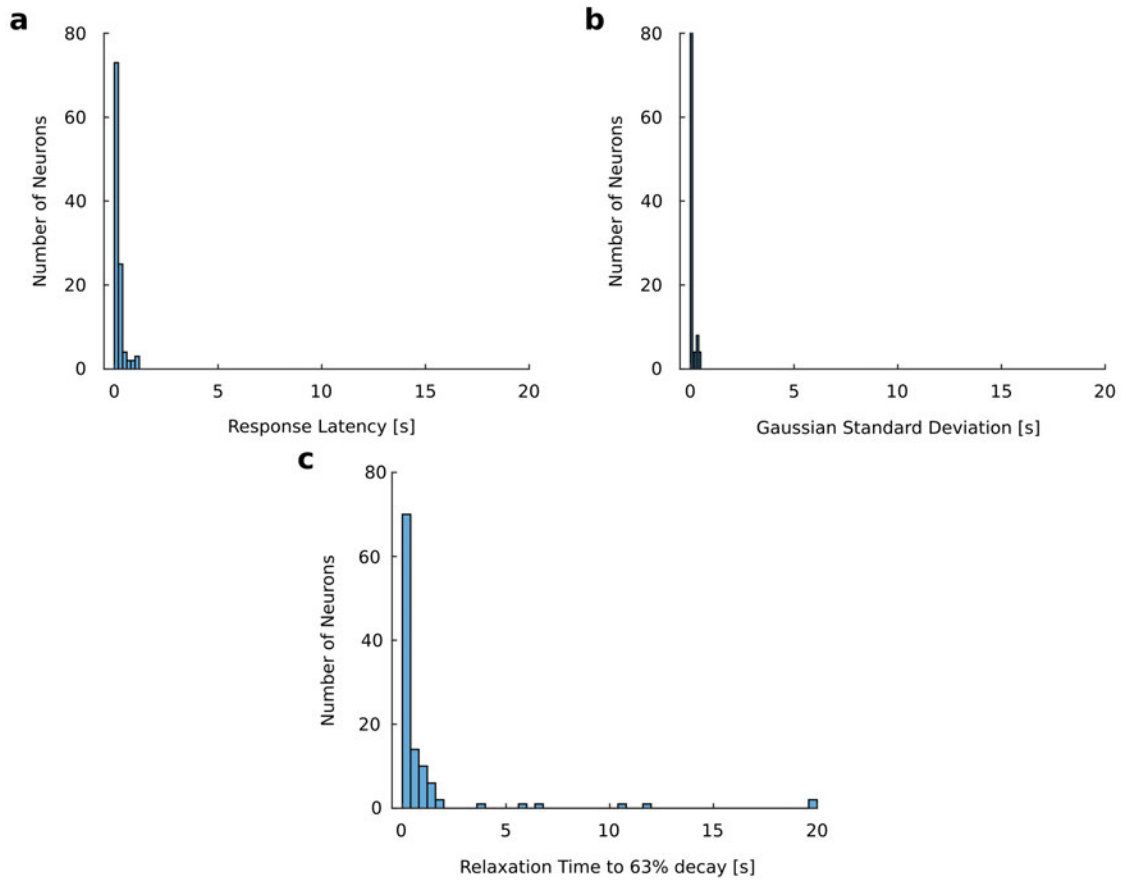


Figure 2-5: The distribution of individual parameters is consistent with temporal context cells. **a**, Neurons started responding (Response Latency) shortly after image onset, and latencies did not cover the entire 5 s viewing period. **b**, Neurons did not show much deviation in their response latencies, with the majority having values of a few hundredths of a second. **c**, Most neurons decayed 63% of the way back to baseline firing (Relaxation Time) quickly, but some showed longer relaxation times that spanned the entire 5 s.

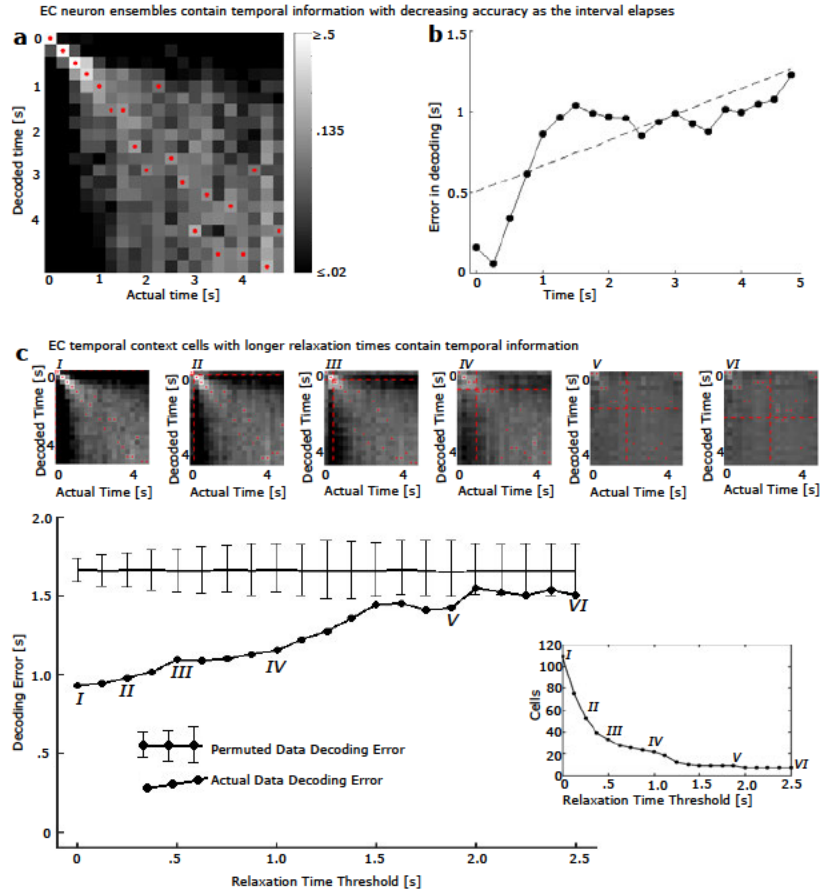


Figure 2-6: The population of entorhinal neurons encodes time with decreasing accuracy as the interval elapses. **a**, Performance of a linear discriminant analysis decoder (LDA) trained on the entire population of EC neurons. The x-axis indicates the actual time bin and the y-axis indicates the decoded time, where lighter shading corresponds to higher probabilities. The most frequently decoded time for each actual time bin is marked with a red dot. **b**, Decoding error increases with the passage of time. The x-axis gives the actual time since image onset; the y-axis gives the mean of the absolute error produced by the decoder in **a**. The dotted black line is a fitted regression line. **c**, The temporal code was distributed along many temporal context cells, including those with slow relaxation times. We decoded time from progressively smaller numbers of temporal context cells, by increasing the smallest relaxation time considered. (Bottom left), the absolute error over all time bins as a function of Relaxation Time Threshold. The entire posterior distribution for selected points labeled by Roman numerals are shown along the top of the figure. The threshold is shown as dashed red lines. The horizontal line shows the absolute error in decoding from a permuted dataset; error bars show the 0.025 and 0.975 quantiles. Right: the number of temporal context cells remaining in the analysis as a function of Relaxation Time Threshold.

2.2.2 Populations of entorhinal temporal context cells carry graded information about time

It is well-understood that hippocampal time cells can be used to decode the time since the beginning of a time interval (Mau et al., 2018). To assess the temporal information present in the population of entorhinal neurons, with special attention to the population of temporal context cells identified by the model-based analysis above, we trained a linear discriminant analysis (LDA) decoder to estimate time following presentation of an image. To the extent the predicted time bin for out-of-sample data is close to the actual time bin, one can conclude that the population response carried information about time. We first describe results from the entire population and then focus on the subpopulation of temporal context cells.

Time was decoded better than chance from the population of entorhinal neurons

Figure 2-6 shows the results of the LDA on all 349 neurons from monkey EC. Our first question was whether or not the population contains information about time. For each time bin in Figure 2-6a, the confidence of the decoder (the posterior distribution) is shown across the range of possible time estimates. Perfect prediction would correspond to a bright diagonal; random decoding would correspond to a uniform gray square. Qualitatively, the non-uniformity of Figure 2-6a suggests that elapsed time can be decoded from the population of EC neurons. To quantitatively assess this, we found that the posterior distribution from the test data was reliably different from a uniform distribution using a chi-squared goodness of fit test, $\chi^2(380) = 3906.8$, $p < 0.0001$.

Supporting this result, the mean absolute value of decoding error from the cross-validated LDA was reliably lower than the decoding error from training with a permuted data set. In each of 1,000 permutations we randomly reassigned the time

bin labels of the training events used to train the classifier. The absolute value of the decoding error for the original data was 0.923 s, which was more accurate than the mean absolute value of the decoding error for all 1,000 permutations. As shown in Figure 2-7, the values for the permuted data were approximately normal with mean 1.65 s and standard deviation 0.04 s, resulting in a z-score of more than 18 ($z = 18.175$). These analyses demonstrate that time since image presentation could be decoded from populations of neurons in monkey EC.

The precision of the time estimate decreased as the interval unfolded

Although the population response in entorhinal cortex could be used to reconstruct time, inspection of Figure 2-6a suggests that the precision of this reconstruction was not constant throughout the interval. Figure 2-6b shows the average absolute value of the decoding error at each time bin. These data suggest that this error increased as a function of time. A linear regression of decoding error as a function of time showed a reliable slope, 0.16 ± 0.03 , as well as intercept 0.55 ± 0.1 , both $p < 0.001$, $R^2 = 0.56$, $df = 18$. The information in the entorhinal population about the time of image presentation decreases in accuracy as the image presentation recedes into the past.

Time can be decoded well past the peak firing of temporal context cells

Theories that proposed the existence of temporal context cells argue that they convey information about time *via* their gradual decay. Another possibility is that the temporal context cells only carry decodable information about time because of their rapid deflection near time zero. If that is the case, then the population of entorhinal neurons should only carry information about time in the period of time close to the Response Time. To assess how far into the interval time could be reconstructed, we repeated the LDA analysis excluding progressively more time bins starting from zero.

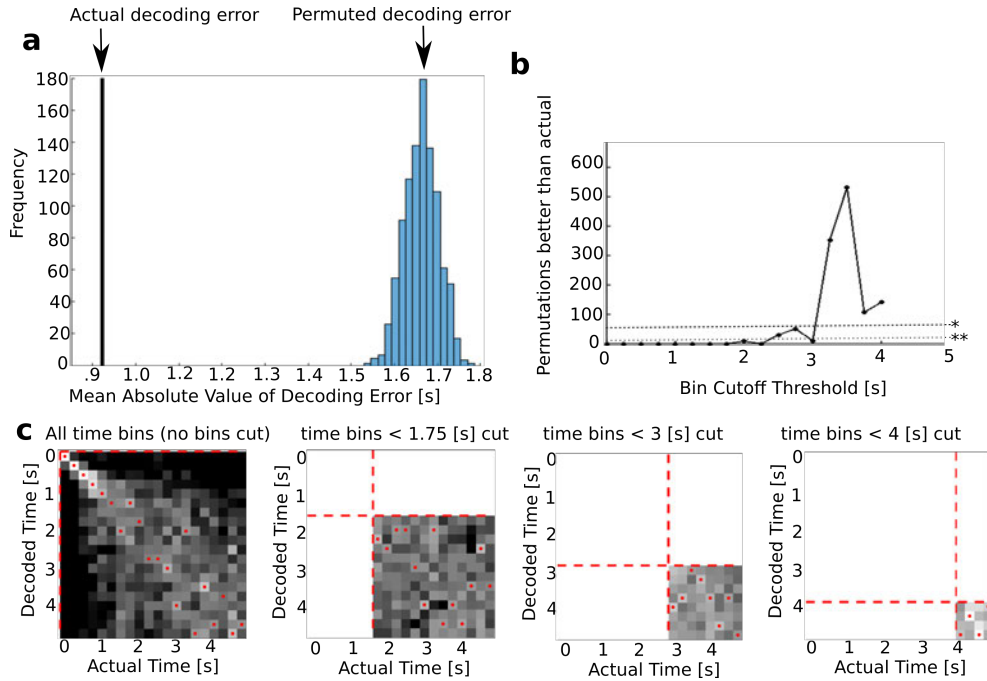


Figure 2-7: Time could be decoded above chance from entorhinal firing. **a**, A histogram of mean absolute value of decoding error for the permuted data compared to actual decoding error. Training data were the same as the actual decoder, except the true time labels were randomly shuffled. The mean absolute value of decoding error for the original data (.93 s) is marked by a vertical line. **b** A plot of decoder performance compared to permuted data as a function of relaxation time threshold. With 1000 permutations, the significance level is the number of permutations that matched or outperformed the actual data divided by 1000 (the line marked by ‘**’ is significant at $p < 0.01$, the line marked by ‘*’ is significant at $p < 0.05$). Thus the decoder is significantly better than chance at $p < 0.01$ up until 2.25 seconds. The decoder is last significantly better than chance at $p < 0.01$ at 3 seconds. This suggests that temporal information is present out to at least 3 seconds. **c** Posterior distributions as bins are dropped. The leftmost sub-panel corresponds to no bins dropped, the next sub-panel corresponds to the first 1.75 seconds dropped the next sub-panel correspond to the first 2.5 seconds dropped, and the rightmost sub-panel corresponds to the first 3.75 seconds dropped) Colorbar and scaling are as in Figure 3a (maximum white correspond to probability 0.5 and greater, minimum black correspond to probability 0.02 and less, color is on log scale).

If the LDA can reconstruct time above chance using only bins corresponding to times $\geq t$, then we can conservatively conclude that time can be reconstructed at least time t into the interval. To assess this quantitatively, the actual data were compared with permuted data for each repetition of the LDA using absolute error to assess performance (see Methods for details, Figure 2.7). This analysis showed that time more than 2.25 s after the image onset can be reliably decoded ($p < 0.01$). This conservative estimate is an order of magnitude longer than the median value of the peak time (0.160 s), suggesting that the gradual decay of temporal context cells could be used to reconstruct information about time. Decoder performance varies later into the delay, with the performance of the decoder actually improving as noisier bins are eliminated from the analysis. For instance, time at 3 s can be reliably decoded ($p < .01$). The decoder is last significantly better than chance at $p < 0.01$ with 3.0 seconds excluded (below the line marked with ‘**’).

Information about time was distributed throughout the population of temporal context cells

Theories that proposed the existence of temporal context cells argue that they convey information about time *via* their gradual decay. To determine how temporal information was distributed across the population of temporal context cells we performed a decoding analysis restricting our attention to temporal context cells. The analysis was performed initially using all temporal context cells, and then progressively removed cells with Relaxation Times shorter than a Relaxation Time Threshold. The Relaxation Time Threshold ranged from 0 s—including all 109 temporal context cells—to 2.5 s—at which point only 7 temporal context cells remained in the analysis. For each Relaxation Time Threshold, performance was summarized by averaging decoding error across all bins. If only a subpopulation of temporal context cells with fast Relaxation Times contributed to the temporal information in the ensemble, we would

expect an abrupt decrease in performance as the Relaxation Time Threshold passed through that critical value.

These results are shown in Figure 2.6c, with selected decoder posteriors shown for various Relaxation Time Thresholds (Figure 2.6cI-VI). The population at each Relaxation Time Threshold was used to generate its own permuted data set. The first observation is that the performance of the decoder changed gradually, suggesting that temporal context cells with a range of Relaxation Times conveyed useful information about the time of image presentation. Examination of the heatmaps in Figure 2.6c suggests that excluding temporal cells with Relaxation Times below a particular value (indicated by dashed red line) disrupts the ability to distinguish times below that value. However, the ability to decode time above the threshold is relatively intact. Statistically, the decoder performed better than chance even with a Relaxation Time Threshold of 1.875 s (with 9 cells remaining). This analysis suggests that temporal information is distributed throughout the population of temporal context cells. Further the population conveys information about a range of time scales because the population has a variety of Relaxation Times.

2.2.3 EC neurons conveyed information about image identity

In this experiment, each image was presented twice. Although it was not practical to assess image coding using a classifier, it was possible to exploit the repetition of images to determine whether EC neurons contained information about image identity. This question was addressed using both single cell analyses and population analyses, which showed convergent results. In both cases we compare the first and second presentation of the same image to the first and second presentations of different images. Note that because these analyses always compare a first presentation to a second presentation, they are not confounded by repetition effects that have been observed in entorhinal neurons (Xiang & Brown, 1998; Meyer & Rust, 2018; Jutras & Buffalo, 2010).

Firing rate of individual neurons was correlated for same presentation of images

For each neuron we assembled an array giving the firing rate during the first presentation of each image (averaged over 5 s) and asked whether this array was correlated with the firing rate of second presentations of the same images. If the firing rate of a neuron depends on the identity of the image, we would expect to see a positive correlation using this measure. For this analysis we restricted our attention to the neurons ($n = 270$) in the entorhinal population that were recorded long enough to be observed for both first and second presentations of a block of stimuli (repetitions were separated by 20-40 minutes).

The mean correlation coefficient (Kendall's τ) across neurons was significantly greater than zero, $\tau = 0.06 \pm 0.02, t(269) = 7.69, p < 0.001$, Cohen's $d = 0.47$, indicating that the spiking activity of many neurons depended on image identity (Figure 2-8b). This comparison was confirmed by a Wilcoxon signed rank test on the values of Kendall's τ , $V = 27577, p < 0.001$. This finding was also observed for the subset of visually responsive neurons ($n = 93$) that we describe as temporal context cells. Taken in isolation, the temporal context neurons showed a mean correlation coefficient significantly greater than zero, as measured by t-test, $0.09 \pm 0.02, t(89) = 7.34, p < 0.001$, Cohen's $d = 0.77$ and Wilcoxon signed rank test, $V = 3553, p < 0.001$. Neurons that were not temporal context cells ($n = 179$) also had a mean correlation coefficient different from zero, $0.04 \pm 0.02, t(176) = 4.65, p < 0.001$, Cohen's $d = 0.35, V = 11363, p < 0.001$. These results are consistent with a population that contains information about stimulus identity.

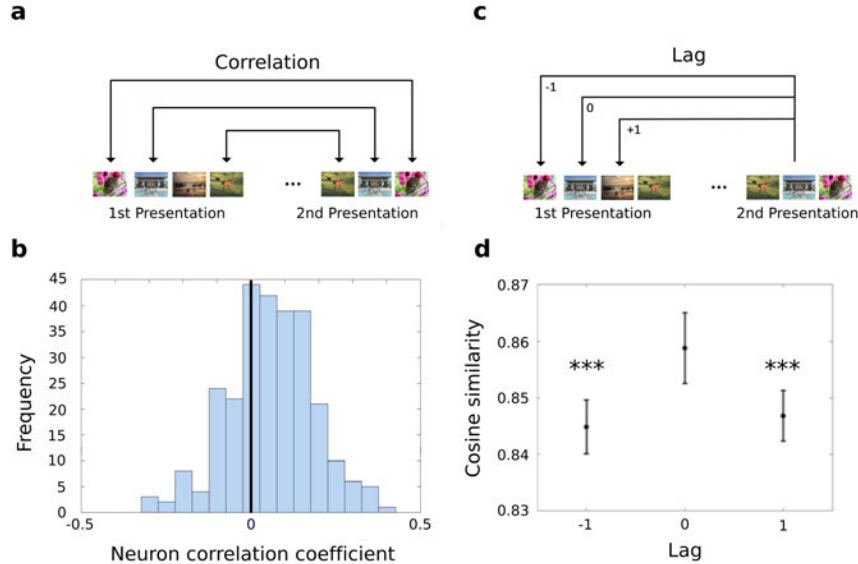


Figure 2-8: The entorhinal ensemble carries information about stimulus identity. **a**, We measured the firing rate of each unit to first and second presentations of each image. For each neuron that was recorded for an entire block of images ($n = 270$), we measured the correlation in firing rate across images using Kendall's τ . **b**, The distribution of Kendall's τ for all entorhinal neurons is shown. This distribution is reliably different from zero ($p < 0.001$, see text for details) indicating that entorhinal firing was sensitive to image identity. **c**, Schematic of a population similarity analysis. We measured the similarity of population response to the second presentation of each image to the first presentation of the same image (lag 0). As controls we also computed the similarity between the population response to the second presentation of an image and the responses to the images neighboring the first presentation of that image. Lag -1 refers to the similarity to the immediate predecessor of the image; lag $+1$ refers to the similarity to the immediate successor of the image. **d**, Cosine similarity of the population response to the second presentation of an image to the original presentation of an image (lag 0) and the images neighboring the original presentation of the image (lags -1 and $+1$) over 64 blocks of first and second presentations. The similarity to lag 0 is greater than the similarity to either of the neighboring images. *** indicates significance at the $p < 0.001$ level. Error bars correspond to the 95 % confidence interval of mean cosine similarity calculated across sessions.

The population of EC neurons was more similar for repeated presentations of the same image

The preceding analyses show that the firing of many EC neurons distinguished image identity above chance. If the response of the entire population contained information about stimulus identity, we would expect, all things equal, that pairs of population vectors corresponding to presentations of the same image would be more similar to one another than pairs of population vectors corresponding to presentations of different images. To control for any potential repetition effect, these analyses compared the similarity between the repetition of an image with its original presentation to the similarity between the second presentation of an image with the first presentation of a different image. To control for any potential recency effects, we swapped images adjacent to the original presentation of the target image. To be more concrete, if we label a sequence of images initially presented in sequence as A, B, and C, we would separately compare the population response to the repetition of B to the response to the initial presentation of A, B, and C. We refer to the similarity of the second presentation of B to the first presentation of B as lag 0. The similarity of the second presentation of B to the first presentation of A is referred to as lag -1 ; the similarity to the first presentation of C is lag $+1$. To the extent that the similarity at lag 0 is greater than lag -1 and $+1$, we can conclude that the population vector distinguishes image identity.

Figure 2-8d shows the results of this population analysis. The similarity for lag 0 pairs (comparing population response to an image with its repetition) was greater than the similarity for lags ± 1 (comparing the response to neighbors of its original presentation). Statistical comparisons to lags ± 1 each showed a reliable difference. A paired t-test comparing population similarity at the level of blocks ($n = 64$ trial blocks of repeated images) showed that similarity at lag 0 was reliably larger than

both lag +1, 0.012 ± 0.005 , $t(63) = 5.11$, $p < 0.001$, Cohen’s $d = 0.64$, and lag -1 , 0.014 ± 0.006 , $t(63) = 5.02$, $p < 0.001$, Cohen’s $d = 0.63$. To evaluate the same hypothesis using a non-parametric method, we performed a permutation analysis by randomly swapping within-session pairs of lag 0 and lag ± 1 and calculating the mean difference between the pairs 100,000 times. The observed value exceeded the value of 100,000/100,000 permuted values for both lags +1 and -1 . We conclude that the population response was more similar for presentations of the same image than for presentations of different images. This analysis, which controlled for repetition and recency, demonstrates that the response of the population of EC neurons reflected image identity. Coupled with the other results reported here, this means that the population carried information about what happened when.

2.3 Discussion

Episodic memory requires information about both the content of an event as well as its temporal context (Tulving, 1983; Eichenbaum et al., 2007; Eichenbaum, 2017). In this study, many EC neurons responded to the onset of the image. These *temporal context cells* responded to image onset at about the same time, within about 300 ms of image onset. However, different temporal context cells showed variable rates of relaxation back to baseline (Figure 2·3). Information about time since the image was presented could be decoded due to gradually decaying firing rates over a few seconds (Figure 2·3d-e). Notably, the relaxation rate was not constant across neurons, but rather showed a spectrum of time constants. The population vectors following repeated presentations of the same image were more similar to one another than to presentations of different images. This, coupled with several control analyses, show that the firing of entorhinal neurons also distinguished stimulus identity. Taken together, the results demonstrate that in the time after image presentation,

the population of EC neurons contained information about what happened when.

Sequentially activated time cells, such as have been observed in the hippocampus (Pastalkova et al., 2008; MacDonald et al., 2011; Salz et al., 2016), medial entorhinal cortex (Kraus et al., 2015) and many other brain regions (Jin et al., 2009; Mello et al., 2015; Tiganj et al., 2018; Tiganj, Shankar, & Howard, 2017) also contain information about what happened when in the past. However, entorhinal temporal context cells have very different firing properties than sequentially activated time cells. As a population, time cells convey the amount of time that has passed since the occurrence of an event by firing at different temporal delays after the triggering event. In contrast, EC temporal context cells all responded at about the same time but relaxed at different rates. The range of relaxation times enables the population to convey information about different time scales. For instance, a temporal context cell that returns to baseline firing within 1 second would not be effective in distinguishing a 5 second interval from a 10 second interval. In contrast, a cell that decays back to baseline around 7 seconds would be effective in distinguishing this longer time period. In this way, a range of decay rates enables the population of temporal context cells to decode time over a wide range of time scales.

2.3.1 Relationship to findings from rodent and monkey EC

The pattern of results observed here aligns well with a recent report from rodent lateral EC (Tsao et al., 2018). In that study, lateral EC neurons changed their firing in response to a salient event, i.e., the animal entering a new environment, and then relaxed back to baseline monotonically. Notably, different neurons relaxed at different rates with time constants ranging from tens of seconds to many minutes. However, despite the many methodological differences between that study and this one—rats moving through a series of open enclosures *vs* seated monkeys observing a series of images—the response properties shared striking similarities, suggesting a common

computational function for EC across species.

The results in this paper are consistent with studies that have shown long-lasting responses in EC neurons in different preparations in rodents. For example, sustained responses have been observed in both in-vitro (Klink & Alonso, 1997; Egorov, Hamam, Fransén, Hasselmo, & Alonso, 2002; Tahvildari, Fransén, Alonso, & Hasselmo, 2007; Yoshida, Fransén, & Hasselmo, 2008; Hyde & Strowbridge, 2012) and anesthetized (Hahn, McFarland, Berberich, Sakmann, & Mehta, 2012; Leitner et al., 2016) approaches. Similarly, slow changes in firing rate were observed in the EC of rats during trace eyeblink conditioning across distinct environmental contexts (Pilkiw et al., 2017). Computational modeling studies have suggested that properties of a calcium non-specific cation current observed in slice are sufficient to generate a spectrum of response decay periods, ranging from brief to prolonged (Tiganj, Hasselmo, & Howard, 2015; Liu et al., 2019). Juxtaposed with work showing spatial responses in navigating rodents (Deshmukh & Knierim, 2011; Wang et al., 2018; Høydal, Skytøen, Andersson, Moser, & Moser, 2019), these findings suggest EC neurons code for “position” in both temporal and spatial domains.

The current results are also consistent with previous findings from the monkey medial temporal lobe. A recent investigation of single-neuron activity across shorter timescales (within a ~ 1 s delay) identified time-varying responses in the hippocampus, but not the entorhinal cortex (Naya et al., 2017). However, in earlier work with longer trial durations, entorhinal neurons exhibited response dynamics across several seconds (Suzuki, Miller, & Desimone, 1997). The current results also mirror previous studies showing coding for temporal information in monkey PFC attributable to slow ramping activity (Machens, Romo, & Brody, 2010; Rossi-Pool et al., 2016, 2019). It remains to be seen if those findings reflect similar or distinct computational mechanisms to those observed in EC in this study.

2.3.2 Exponentially-decaying neurons with a spectrum of time constants is the Laplace transform of time

Why would the brain use two distinct coding schemes—time cells versus temporal context cells—to represent time? One proposed answer is that there might be a local circuit processing advantage from having time cells that can signal a specific moment at the single-neuron level, instead of having that information exist in the collective responses of different temporal context cells. Mechanistically, the brain may achieve creating a time cell response by combining the responses of temporal context cells (Shankar & Howard, 2012). The mathematical description of this process would begin with the brain estimating a temporal record of the past as neural activity that is the real Laplace transform of a function of past time (Shankar & Howard, 2012, 2013; Howard et al., 2014).

Cells coding for the real Laplace transform have exponential receptive fields with a variety of rate constants, very much like the results observed here in entorhinal neurons (Figure 2·3). In this proposal, the activity representing the contents of the past changes in the time after an image presentation. At a time t after image presentation, the neural representation of the past is a function with the presentation of the image at time t . As t increases, this function changes smoothly. A population of neurons coding the real Laplace transform of time should thus change shortly after image presentation and then relax exponentially, with different neurons relaxing at different rates.

Time cells like those in the hippocampus, in contrast, can be described by performing an additional computation upon the activity described above. The inverse Laplace transform — which can be approximated using a set of feed-forward connections with center-surround weights on the activity described above — directly estimates what happened when. Instead of exponential receptive fields, cells coding for the inverse

transform have circumscribed receptive fields that tile the time axis. As the image presentation recedes into the past, neural response to its presentation resembling the inverse transform would generate a series of sequentially-activated time cells.

Laplace transforms of other variables in the MTL

This computational framework for representing functions over continuous variables using the Laplace transform and its inverse can be generalized from time to other variables as well (Howard et al., 2014). For instance, border cells in EC (Boccaro et al., 2010) code for distance from an environmental landmark with monotonically-decaying receptive fields. If the firing profile of border cells is exponential, and if the parameter controlling the spatial sensitivity of this profile differs across neurons, analogous to the differing relaxation times in Figure 2, then a population of border cells would code for the real Laplace transform of distance to the border. Applying the inverse transform to such a population would generate boundary vector cells (Lever, Burton, Jeewajee, O’Keefe, & Burgess, 2009), which have been observed in the subiculum and have been argued to drive classic hippocampal place fields (Burgess & O’Keefe, 1996). Analogously, trajectory coding cells and splitter cells observed in the EC and hippocampus can be understood as coding for the Laplace transform and inverse of functions over ordinal position—the sequence of movements leading up to the present (Frank, Brown, & Wilson, 2000; Wood, Dudchenko, Robitsek, & Eichenbaum, 2000). More broadly, this computational framework can be used to generate functions over a variety of spatiotemporal trajectories, which has been proposed to be a basic function the the MTL (Hasselmo, 2012; Dannenberg, Kelley, Hoyland, Monaghan, & Hasselmo, 2019). In this view, time cells and place cells are just two cases of a more general computational function (Kraus et al., 2013, 2015; Howard & Eichenbaum, 2015).

Laplace transforms of time throughout the brain

If the brain contains a compressed record of the past (James, 1890; Husserl, 1966) as a neural representation across many different “kinds” of memory (Chater & Brown, 2008; Howard et al., 2015), then one might expect the existence of neurons with conjunctive receptive fields for what happened when across many different brain regions. Indeed, stimulus-specific time cells coding for what happened when have not only been found in regions believed to be important for episodic memory (MacDonald et al., 2013; Taxidis et al., 2018), but also regions that support working memory (Tiganj et al., 2018; Cruzado et al., 2020) and classical conditioning (Adler et al., 2012).

Because computational work using the Laplace transform has shown that a population of time cells can be constructed from temporal context cells (Shankar & Howard, 2012, 2013; Howard et al., 2014), we may speculate about the pervasiveness of this phenomenon across the brain. Perhaps temporal context cells may be found in other brain regions outside the EC to support this computation across separate regions of the brain. Alternatively, especially given the high sensory convergence within the EC, perhaps temporal context cells in the EC are utilized for generating time cells across the brain.

2.4 Methods

2.4.1 Subjects, training, and surgery

Two, male rhesus macaques (*Macaca mulatta*), 10 and 11 years old, and weighing 13.8 and 16.7 kg respectively, were trained to sit in a primate chair (Crist Instrument Company, Inc., Hagerstown, MD) and to release a touch-bar for fruit slurry reward delivered through a tube. The monkeys were trained to perform various tasks by releasing the touch-bar at appropriate times relative to visual stimuli presented on a screen. Magnetic resonance images of each monkey’s head were made both before and

after surgery to plan and confirm implant placement. Separate surgeries were performed to implant a head post, then months later, a recording chamber, and finally a craniotomy within the chamber. All experiments were performed in accordance with protocols approved by the Emory University and University of Washington Institutional Animal Care and Use Committees.

2.4.2 Electrophysiology

Each recording session, a laminar electrode array (AXIAL array with 13 channels, FHC, Inc.) mounted on a microdrive (FHC, Inc.) was slowly lowered into the brain through the craniotomy. Magnetic resonant images along with the neural signal were used to guide the penetration. Spikes and local field potentials were recorded using hardware and software from Blackrock, Inc., and neural data were sampled at 30 kHz. A 500 Hz high-pass filter was applied, as well as an electric line cancellation at 60 Hz. Spikes were sorted offline into distinct clusters using principal components analysis (Offline Sorter, Plexon, Inc.). Sorted clusters were then processed further by custom code in MATLAB to eliminate any data where minimum inter-spike interval was less than 0.001 s, and to identify any missed changes in signal (e.g., decreased amplitude in the waveform of interest, a new waveform appearing), using raster plots and plots of waveforms across the session for each neuron. When change in signal was identified, appropriate cuts were made to exclude compromised spike data from before or after a change point. 455 potential single neurons originally isolated in Offline Sorter were reduced to 357 single neurons. To further ensure recording location within the entorhinal cortex and identify from which cortical layers neurons were recorded, we examined each session's data for the stereotypical, electrophysiological signature produced across entorhinal cortical layers at the onset of saccadic eye movement (Killian, Jutras, & Buffalo, 2012; Killian, Potter, & Buffalo, 2015; Meister & Buffalo, 2018). Recording sessions took place in both anterior and posterior regions of the

entorhinal cortex. One recording session, which other electrode placement metrics suggest was conducted above the entorhinal cortex within the hippocampus, lacked this electrophysiological signature and was excluded from further analysis (8 single neurons were excluded from this session). No recording sessions showed the current source density electrophysiological signature of adjacent perirhinal cortex (Takeuchi, Hirabayashi, Tamura, & Miyashita, 2011) at stimulus onset.

2.4.3 Experimental design and behavioral task

For all recordings, the monkey was seated in a dark room, head fixed and positioned so that the center of the screen (54.1 cm \times 29.9 cm LCD screen, 120 Hz refresh rate, 1280 \times 720 pixels, BenQ America Corp., Costa Mesa, CA) was aligned with his neutral gaze position and 60 cm away from the plane of the his eyes (equating to 25 screen pixels per degree of visual angle, or 1/cm). Stimulus presentation was controlled by a PC running Cortex software (National Institute of Mental Health, Bethesda, MD). Gaze location was monitored at 240 Hz with an infrared eye-tracking system (I-SCAN, Inc., Woburn, MA). Gaze location was calibrated before and during each recording session with calibration trials in which the monkey held a touch-sensitive bar while fixating a small (0.5) gray square presented at various locations on the screen. The square turned yellow after a brief delay chosen uniformly from the interval from 0.40 s to 0.75 s. The monkey was required to release the bar in response to the color change for delivery of the fruit slurry reward. The subtlety of the color change forced the monkey to fixate the location of the small square to correctly perform those trials, therefore allowing calibration of gaze position to the displayed stimuli. Specifically, the gain and offset of the recorded gaze position were adjusted so that gaze position matched the position of the fixated stimulus. Throughout the session, intermittent calibration trials enabled continual monitoring of the quality of gaze position data and correction of any drift.

Before each image presentation, a crosshair (0.3×0.3) appeared in one of eighteen possible screen locations. Once gaze position registered within a 3×3 window around the crosshair and was maintained within that spatial window for between 0.50 and 0.75 s (chosen uniformly), the image was presented. Images were large, complex images downloaded from the public photo-sharing website, Flickr (www.flickr.com). If necessary, images were re-sized by the experimenter for stimulus presentation (sized 30×15 for Monkey WR and 30×25 for Monkey MP). Monkeys freely viewed the image, and then the image vanished after gaze position had registered within the image frame for a cumulative 5 seconds. No food reward was given during image viewing trials. Each image presentation was followed by three calibration trials.

Image stimuli were unique to each session, and each image was presented twice within a session about 20 to 40 minutes apart. Images were presented in a block design so that novel and previously-seen images were presented throughout the session. Within a trial block, novel images (30 or 60) would first be shown, and then presented again in pseudorandom order. After completing a block of trials, a new block of trials would begin. In the first 16 sessions, a three-block design of 60 image presentations (30 novel) per block was used, with a total maximum of 180 image presentations per session. In the rest of the sessions ($n = 41$), there were a total maximum of 240 image presentations across two trial blocks (120 image presentations of which 60 were novel within each trial block).

2.4.4 Analysis of Neural Firing Fields

In order to determine temporal firing fields, spikes were analyzed using a custom maximum likelihood estimation script run in MATLAB 2016a. We calculated model fits across all trials available for each particular neuron considering the time from 500 ms before image presentation to 5 s after image presentation. Fits of nested models were compared using a likelihood ratio test. In the present paper, we considered three

models: a constant firing model, a model adding a Gaussian time term, and an ex-Gaussian model for which the time term was given by the convolution of a Gaussian and an exponential time term. The constant model,

$$M_{\text{const}}(t; a_o) = a_o \quad (2.1)$$

consisted of a single parameter a_o that predicted the constant probability of a spike at each time t .

The ex-Gaussian model describes the temporally-modulation of the firing field as the convolution of a Gaussian function with an exponentially decaying function:

$$M_{\text{ex-gauss}}(t; a_o, a_1, \sigma, \mu, \tau) = a_o + a_1 \int_{-\infty}^{\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} e^{-\frac{t}{\tau}} dt \quad (2.2)$$

The ex-Gaussian distribution has been used extensively in studies of human response time data for many years (Ratcliff & Murdock, 1976). In the limit as $\tau \rightarrow 0$ the exponential function becomes a delta function and the result of the convolution in Eq. 2.2 is a Gaussian function. Similarly, in the limit as $\sigma \rightarrow 0$ the Gaussian function becomes a delta function and the result of the convolution is an exponential function starting at μ . As such, this model is able to describe a range of peak firing times as well as varying degrees of skew (Fig. 2.1).

Two terms, a_o and a_1 describe the contributions of the constant and time-modulated terms. Three parameters describe the shape of the temporally-modulated term (Figure S1). μ and σ describe the mean and standard deviation of the Gaussian distribution, which estimates the time that a neuron's response maximally deviates from baseline, and the variability in that response time respectively. In the text, μ is referred to as the Response Time. τ measures the time constant of the exponential decay, and captures the time that a neuron has returned 63% of the way back to baseline. In the text, τ is referred to as the Relaxation Time.

To estimate parameters of Eq. 2.2 numerically we used an explicit form for the solution of the convolution in Eq. 2.2:

$$M_{\text{EG}}(t; a_o, a_1, \sigma, \mu, \tau) = a_o + \frac{a_1}{2} e^{\frac{(2\mu + \frac{\sigma^2}{\tau} - 2t)}{2\tau}} \operatorname{erfc}\left(\frac{\mu + \frac{\sigma^2}{\tau} - t}{\sqrt{2}\sigma}\right) \quad (2.3)$$

where erfc is the complementary error function. μ was allowed to take values between 0 and 5 s. τ was allowed to take values between 0 and 20 s. σ was allowed to take values between 0 and 1 s. Likelihood was estimated in a 5.5 s long window that included the 0.5 s prior to presentation of the image and the 5 s after presentation of the image.

We evaluated the models for each neuron *via* a likelihood ratio test and counted the number of neurons that were 1) better fit by the ex-Gaussian model at the 0.05 level, Bonferonni-corrected by the total number of 349 neurons, 2) changed their firing by at least 1 Hz, and 3) reached a firing rate of at least 3 Hz. In addition, we required that both even and odd trials for a neuron were significantly fit by the model, and that the those fits had a Pearson's correlation coefficient greater than 0.4.

2.4.5 Linear Discriminant Analysis (LDA)

An LDA classifier was used to decode time since onset of the image from the population using information from all neurons.

LDA Implementation

Even and odd trials were used for training and testing respectively. The number of available trials varied for each neuron. To mitigate any problems from this, several steps were taken. First, four neurons with less than 30 trials each were entirely excluded from this analysis. Neurons with less than 200 trials were bootstrapped to 200 trials, while neurons with more than 200 trials were randomly down-sampled. Time was discretized into 0.25 s bins. For each bin of each trial, the firing rate was

calculated across neurons. To avoid errors due to a singular covariance matrix, a small amount of uniform noise (between 0 and 1×10^{-13} Hz) was added to the firing rate in each time bin. The averaged firing rate of each time bin for each training trial across all neurons made up an element of the training data. The averaged firing rate of each time bin for each testing trial across all neurons made up an element of the testing data. LDA was implemented using the MATLAB function “classify.” This function takes in the training data, testing data, labels for the training data, and a selection of the method of estimation for the covariance matrix (the option “linear” was used) and returns a posterior distribution across bins for each test trial.

Estimating the duration of temporal coding

To assess the quality of temporal information at different points within the interval, the LDA was repeated for successively fewer bins, at each step removing the earliest time bin. If time since presentation of the image can be decoded above chance using only information after time T , one can conclude that the population contained temporal information about time at least a time T after presentation of the image. For each repetition the decoder was tested by training it on data with permuted time labels. We compared the absolute error of the actual data to the distribution generated from 1,000 permutations. The classifier’s performance was considered significantly better than chance if fewer than 10/1,000 permutations gave a better result than the unpermuted data (corresponding roughly to $p < .01$).

Evaluating the distribution of temporal information over temporal context cells with different Relaxation Times

To assess the distribution of temporal information over temporal context cells as a function of their Relaxation Time, neurons with small Relaxation Times were progressively omitted from the LDA. First all temporal context cells were used (corre-

sponding to a Relaxation Time Threshold of zero), then only cells with Relaxation Time longer than 0.125 seconds, then only cells with Relaxation Time longer than 0.25 seconds, and so forth. The longest Relaxation Time Threshold evaluated was 2.5 seconds. Performance was parametrized by averaging the absolute value of decoding error across all time bins. As a control, for each pseudo sub-population of cell, bins with permuted labels were also trained on and decoded from 1000 times.

2.4.6 Stimulus sparsity analysis

To assess stimulus specificity in a way that facilitates comparison to previous human work, we followed the analysis method of (Mormann et al., 2008). In order to determine how many images a given neuron responded to, for each neuron we formed a distribution of baseline firing rates from the 500 ms prior to image onset on each trial. For each image, we took the trials in which they appeared and binned the firing rates of the 1 s following image onset into 19 overlapping bins, each of which were 100 ms long. We then compared these binned firing rates to the baseline distribution via a two-tailed Mann-Whitney U test, using the Simes procedure and a conservative significance threshold of $p = 0.001$. For each neuron, we counted the number of images that exceeded this threshold

2.4.7 Population vector analysis of stimulus specificity

We constructed population vectors to evaluate the degree to which the entire population of entorhinal neurons was sensitive to the identities of visual images. For each repeated image, we created two population vectors, one corresponding to the first presentation and one corresponding to the second presentation. Each vector was created from the mean firing activity of all neurons recorded in a session during the 5 s of free viewing. Mean firing rates were normalized by each neuron’s maximum average firing rate so that firing rates ranged from 0 to 1. Only blocks where all

images were presented twice were considered. In order to control for different block lengths between sessions, only the first 30 images presented in each block were used. All neurons that were recorded during first and second presentations of an image were included in this analysis ($N = 332$). The average number of simultaneously recorded neurons in a block was 8.51, with a standard deviation of 4.04, range 2-19. Similarity was measured by the cosine similarity of the two population vectors. We compared the cosine similarity of two presentations of the same image to the first presentation of one image and the second presentation of a different image. As a control we instead compared the population vector from the repetition of an image to the adjacent near-neighbors of the original image presentation. Near neighbors were required to be the first presentation of an image. Within session error bars represent the 95% confidence interval (Morey, 2008).

Chapter 3

The time to initiate retrieval of a memory depends on recency.

In recognition experiments, participants must judge if a given probe was previously experienced or new. As the recency of a repeated probe decreases, the accuracy of this judgment decreases, and response times (RT) increase (Shepard & Teghtsoonian, 1961; B. B. Murdock & Anderson, 1975; Monsell, 1978; Hockley, 1982; Donkin & Nosofsky, 2012). Despite decades of empirical and modeling work, the cause of this recency effect remains unclear. In continuous recognition, there is no separation between a study phase and a test phase; the participant must make a new/old judgment on every trial. Consider the task of an individual engaged in continuous recognition. The individual must correctly identify an item as old by comparing it to the contents of their memory. In continuous recognition, the recency effect manifests as a sublinear increase in RT with increasing lag of the repeated probe (D. L. Hintzman, 1969; Okada, 1971; Hockley, 1982). Hockley (1982) in particular, found a logarithmic increase in RT with increasing lag. A question that remains however is, why does it take longer to retrieve memories experienced further in the past? One hypothesis is that the strength of memory traces decay over time or with intervening items. A second hypothesis is that it takes longer to retrieve memories from further in the past. These hypotheses are not mutually exclusive.

Strength models are consistent with traditional signal detection models of recognition memory (B. B. Murdock & Dufty, 1972; Wixted, 2007; Rotello, 2017), in which

the output of memory for each probe consists of a scalar decision variable. Strength models can be implemented in distributed memory models that assume that memory is a composite store containing a noisy record of features from all the studied items (e.g., J. A. Anderson, 1973; B. B. Murdock, 1982; Shiffrin, Ratcliff, Murnane, & Nobel, 1993). A composite memory store can account for the recency effect if the features of items experienced further in the past are stored with less fidelity than items experienced more recently. Indeed the confidence/RT relationship (D. A. Norman & Wickelgren, 1969; B. B. Murdock & Dufty, 1972) is consistent with the hypothesis that old probes that evoke more strength should result in both faster RTs and high confidence in their prior occurrence. A strength model can account for the logarithmic relationship between RT and recency increase if the strength of the match between a probe and the contents of memory decreases appropriately (J. R. Anderson, Bothell, Lebiere, & Matessa, 1998) and if this strength is coupled with a model of information accumulation (e.g., S. Brown & Heathcote, 2005; Ratcliff, 1978; Usher & McClelland, 2001). The key feature of strength models is that information about all traces becomes available at the same time during retrieval.

Another class of models in recognition memory hypothesize that at least some correct judgments in item recognition result from a detailed recollective process (Mandler, 1980; Tulving, 1985; Yonelinas, 1997). Recollection results in the availability of detailed source information about the encoding context of the probe stimulus and this retrieval process can succeed or fail (Province & Rouder, 2012; Kellen & Klauer, 2014). Although the output of recollection need not be constant (Onyper, Zhang, & Howard, 2010) the recollective process should take some time before it provides information about the probe (D. L. Hintzman & Curran, 1994). If the time to recover an episodic memory depends on its recency, a discrete state model could also account for the logarithmic effect on RT. It has been proposed that recollection

depends on recovery of a gradually-changing temporal context (Tulving & Madigan, 1970; G. Schwartz et al., 2005; Folkerts et al., 2018). Perhaps the time to retrieve a prior state of temporal context depends on its “distance” from the present state of temporal context. The finding that RT in continuous recognition increases with the log of lag reflects the fact that temporal context changes as a function of log time (G. D. A. Brown, Neath, & Chater, 2007; Howard et al., 2015; Howard, 2018; Cao et al., 2021).

These two hypotheses—the recency effect is due to decaying trace strength (D. A. Norman & Wickelgren, 1969) *vs* the recency effect is due to a discrete retrieval process that takes an amount of time that depends on recency—can be distinguished by examining the shape of the RT distributions (Figure 3-1, top). If recency only affects the strength of a memory trace, then the time needed to begin evaluating a given memory is the same regardless of how far in the past the probe was experienced. The strength of that match, however, should depend on the probe’s recency. If, however, old judgments must await the termination of a discrete retrieval process, then RT distributions should rise at different times (Figure 3-1, bottom). If there is no difference in the quantity of evidence that is retrieved by this discrete retrieval process, then the distributions at different lags will maintain a constant offset from each other as one moves through the distribution. Critically, a change in the time to initiate the search as a function of lag cannot be accounted for by a purely strength based account of the recency effect.

While a shift in the RT distribution is consistent with a recency dependent retrieval process during the memory comparison phase, this is not the only possible explanation. The probe must be encoded before it can be compared to memory. A shift in the RT distributions would also be consistent with the hypothesis that recently-experienced probes are processed faster *prior* to memory retrieval. If the

recency of a repeated item allows it to be processed faster as a probe, then repeating the item again should have an additional effect on RT. This implies that the time it takes for an item to become available in memory should depend on the recency of both previous presentations. If changes are driven by improved processing of recently-presented probes, then the recency of both previous presentations ought to both affect RT. In contrast, if the recency effect instead arises due to a recency dependent delay in initiating retrieval of a memory, then the delay would only depend on the most recent lag. Further, this would predict that the effect of recency should be the same on both the first and second presentations.

To the best of our knowledge, a systematic change in the time to initiate the memory search has not previously been observed in continuous recognition. The main issue in continuous recognition is that as lag increases, accuracy decreases (Hockley, 1982; Shepard & Teghtsoonian, 1961), making it more challenging to measure the effect of recency on retrieval dynamics independently of changes in accuracy. Brady, Konkle, Alvarez, and Oliva (2008) showed participants hundreds of memorable images in a continuous recognition task with lags varying over more than two orders of magnitude, with lags from 1 (no intervening items) to 128. In addition, the RT data from Brady et al. (2008) are minimally affected by sequential dependencies, which are known to affect RTs in recognition memory (Malmberg & Annis, 2012), as repeated items could not take place in adjacent trials. Because of the wide range of lags tested, and the elimination of sequential dependencies, Brady et al. (2008) is well-suited to study the effect of recency on RT distributions. This paper analyzes the RT data collected during the Brady et al. (2008) task (referred to as Experiment 1), and five additional experiments designed to assess the generality of this finding. All six experiments show a systematic change in the rise time of RT distributions as a function of log lag, suggesting a discrete retrieval process that depends on the recency

of the to-be-retrieved memory. Experiment 6 assesses the alternative hypothesis that the change in the rise time of RT distributions is due to probe fluency by repeating old items up to five times, presumably saturating probe fluency. Despite an overall decrease in RT for repeated items—indicating that the manipulation of probe fluency was successful—the effect of recency on RT was unchanged.

3.1 Methods

Table 3.1 contains a summary of the major methodological differences between the six continuous recognition experiments. In short, experiment 2 was a replication of experiment 1 (Brady et al., 2008) using a different subject population (i.e., Amazon M-Turk or Boston University Students). Experiments 3-6 displayed phase scrambled noise rather than a fixation cross between each image. In experiment 4, words from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982) were used as stimuli rather than images. In experiments 1-4, participants were only required to respond if they believed the stimuli had previously been presented, in experiments 5-6 participants had to respond new/old on every trial. Finally in experiments 1-5, some stimuli were repeated a second time while in experiment 6 some stimuli were repeated five times.

3.1.1 Experiment 1

Participants

Fourteen individuals participated in this study in exchange for financial compensation. Participants were recruited via Amazon Mechanical Turk. All participants completed the task simultaneously, and completed the task at computer workstations that were matched for similar screen size and viewing distance.

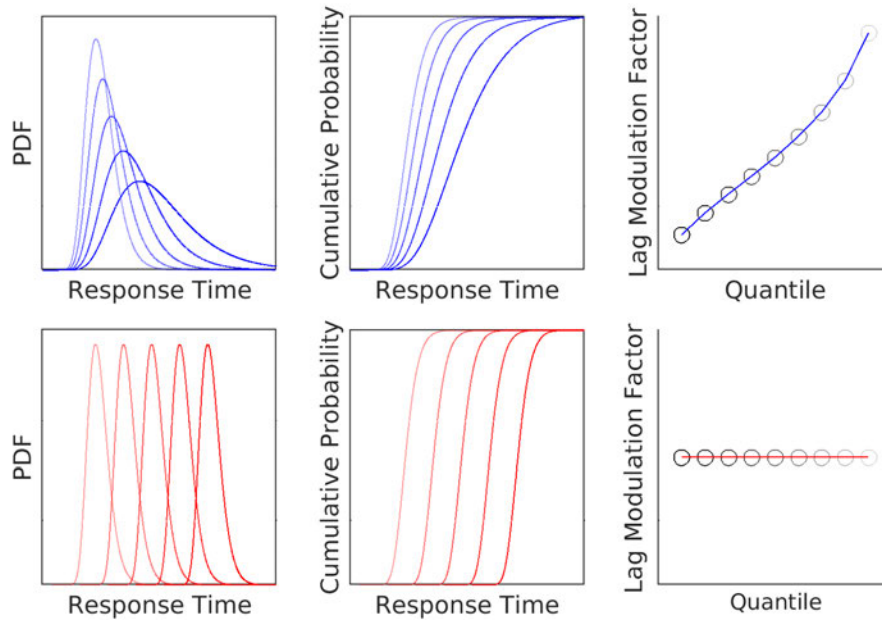


Figure 3-1: Distinguishing two potential accounts for the recency effect from response time distributions. **Top.** In strength-based accounts of recency, all items are available simultaneously but older items (darker lines) are represented with less fidelity than newer items (lighter lines). *Left:* Probability density functions of response times for each lag. The shape of the distribution changes with the drift rate. *Middle:* A cumulative distribution of response times for different lags. RT distributions start at the same point regardless of lag, but spread as you move into later quantiles. *Right:* A plot of lag modulation factor as a function of quantile. In strength-based accounts of recency, lag modulation factor is zero at the start of the distribution and monotonically increases for later quantiles. **Bottom.** If the recency effect arises due to a delay in recovering older temporal contexts, then older items require more time before evidence accumulation can begin. *Left:* Probability density functions of response times for each lag. Rather than varying in their drift rate, newer items begin accumulating evidence before older items. *Middle:* A cumulative distribution of response times for different lags. Because the distributions do not change their shape, the distance between response times is similar across deciles. *Right:* The Lag Modulation Factor as a function of quantile. The change in the start to accumulate evidence results in a non-zero intercept.

	Exp. 1	Exp. 2
Subject Pool	Amazon M-Turk	BU Students
Inter-trial display	Fixation Cross	Fixation Cross
Stimuli	Images	Images
Response Instructions	Old Only	Old Only
Number of Repetitions	2	2
Number of Subjects	14	29
	Exp. 3	Exp. 4
Subject Pool	BU Students	BU Students
Inter-trial display	Phase Scrambled Noise	Phase Scrambled Noise
Stimuli	Images	Words
Response Instructions	Old Only	Old Only
Number of Repetitions	2	2
Number of Subjects	39	33
	Exp. 5	Exp. 6
Subject Pool	BU Students	BU Students
Inter-trial display	Phase Scrambled Noise	Phase Scrambled Noise
Stimuli	Images	Images
Response Instructions	New/Old	New/Old
Number of Repetitions	2	5
Number of Subjects	41	46

Table 3.1: The major methodological differences between the six continuous recognition experiments. The experiments varied on the subject pool that participants were recruited from (Amazon M-Turk or Boston University Students), what was displayed between each trial (fixation cross or phase scrambled noise), the type of stimuli used (images or words), the instructions given to the participants (respond only to repeated items or indicate if the current item is new/old), and the number of times an item could be repeated (2 or 5).

Materials

Stimuli were assembled from a commercially available database (Hemera Photo-Objects, Vol. I and II) and searches done on Google Image Search. Each participant viewed a series of 2500 images, 396 of which were randomly chosen to be repeated. Of these 396 images, 56 were repeated immediately (lag 1), 52 were repeated such that the previous presentation was two trials prior (lag 2), 48 were repeated four trials after their previous presentation, and so on out to 16 images repeated at lag 1024. Lags greater than 128 were not considered, as repetitions occurred across blocks. No two consecutive trials were allowed to contain repeated images, preventing sequential effects. Some images were repeated twice, at lags greater than 1, to avoid adjacent repeat trials. In order to minimize the effect of fatigue, and to simplify counterbalancing when repeats of various lags would occur, the trials with repeated items were the same for all participants. However, all participants viewed images in different orders, and different images were repeated for each participant.

Procedure

Participants viewed ten blocks of images, which lasted around 20 minutes each. All blocks contained 290 trials, except for the first block which consisted of 289 trials, and the tenth and final block which consisted of 287 trials. Participants were given a 5 minute break between each block, and instructed to not discuss any images they had seen. Images were presented for 3 seconds, with a fixation cross presented for 800 ms between each image. The stimuli, approximately 7.5° by 7.5° of visual angle, were presented one at a time in the center of the screen. Participants were instructed to hit the space bar if the current image had previously been presented. They did not need to do anything on trials where they judged the current image to be new. Participants only received feedback on trials in which they responded. On false alarm

trials the fixation cross following the trial was red, and on trials which were hits, the fixation cross turned green.

3.1.2 Experiment 2

Participants

Twenty-nine individuals participated and were compensated with \$15 for their time. All participants were members of the Boston University community, and were recruited via the quick job page. Participants gave informed written consent prior to beginning the study, and the protocol was approved by the Boston University Charles River Institutional Review Board. Participants completed the task at different times at one of two 21.5 inch, mid-2011 iMac computers.

Materials

The experiment was implemented in PyEPL (Geller, Schlefer, Sederberg, Jacobs, & Kahana, 2007). Stimuli were pulled from the same set of images as used in experiment 1. Each participant viewed a series of 650 images, 250 of which were randomly chosen to be repeated. Of those 250 images, 30 were repeated immediately (lag 1), 28 were repeated at a lag of 2, and so on out to 16 images repeated at lag 128. Repetitions that occurred across blocks were not considered. For each subject, 27 items were repeated a second time at lags 2, 4, and 16 (9 per lag). For items repeated a second time, the lag between its first presentation and first repetition was restricted to 8, 16, and 64. The presentation order of images and the trials which contained repetitions were randomized for each subject.

Procedure

Participants viewed two blocks of images, which lasted around 20 minutes each. Both blocks consisted of 450 trials. Participants were given a short break (5 minutes)

between the two blocks. Images were presented for 2.6 seconds, with a fixation cross presented for 400 ms between each image. The stimuli, approximately 7.5° by 7.5° of visual angle, were presented one at a time in the center of the screen. Participants were instructed to hit the space bar if the current image had previously been presented. They did not need to respond on trials where they judged the current image to be new. Participants only received feedback on trials in which they responded. On false alarm trials the fixation cross following the trial was red, and on trials which were hits, the fixation cross turned green.

3.1.3 Experiment 3

Participants

Thirty-nine individuals participated and were compensated with \$15 for their time. All participants were members of the Boston University community, and were recruited via the quick job page. Participants gave informed written consent prior to beginning the study, and the protocol was approved by the Boston University Charles River Institutional Review Board. Participants completed the task at different times at one of two 21.5 inch, mid-2011 iMac computers.

Materials

The experiment was implemented in PyEPL (Geller et al., 2007). Stimuli were assembled from the same set of images as used in the previous experiments. In this experiment, there was a wider range of lag values, ranging from lag 1 to lag 512. In order to better accommodate these longer lags, participants viewed a series of 1360 images, 276 of which were randomly chosen to be repeated. Of those 276 repeated images, 30 were chosen to be repeated immediately (lag 1), 28 were repeated at a lag of 2, and so on out to 12 images repeated at lag 512. Repetitions that occurred across blocks were not considered. For each subject, 27 items were repeated a second time

at lags 2, 4, and 16 (9 per lag). For items repeated a second time, the lag between its first presentation and first repetition was restricted to 8, 16, and 64. As in experiment 2, both the order of the images, and the trials where repeated items occurred were randomized for each participant.

Procedure

Participants viewed two blocks of images, which lasted around 35 minutes each. Both blocks consisted of 680 trials. Participants were given a short break (5 minutes) between the two blocks. Images were presented for 2.6 seconds, however instead of a fixation cross, phase scrambled images were used as masked separators and presented for 400 ms between each image. The stimuli, approximately 7.5° by 7.5° in visual angle, were presented one at a time in the center of the screen. Participants were instructed to hit the space bar if the current image had previously been presented. They did not need to do anything on trials where they judged the current image to be new. Participants only received feedback on trials in which they responded. On false alarm trials a red square appeared around the image and on trials which were hits, a green square appeared around the image.

3.1.4 Experiment 4

Participants

Thirty-three individuals participated and were compensated with \$15 for their time. All participants were members of the Boston University community, and were recruited via the quick job page. Participants gave informed written consent prior to beginning the study, and the protocol was approved by the Boston University Charles River Institutional Review Board. Participants completed the task at different times at one of two 21.5 inch, mid-2011 iMac computers.

Materials

The experiment was implemented in PyEPL (Geller et al., 2007). Stimuli were assembled from the Toronto Word Pool (Friendly et al., 1982), a list of common English nouns. Participants viewed a series of 1360 words, 276 of which were randomly chosen to be repeated. Of those 276 repeated words, 30 were chosen to be repeated immediately (lag 1), 28 were repeated at a lag of 2, and so on out to 12 images repeated at lag 512. Repetitions that occurred across blocks were not considered. For each subject, 27 items were repeated a second time at lags 2, 4, and 16 (9 per lag). For items repeated a second time, the lag between its first presentation and first repetition was restricted to 8, 16, 64. Both the order of the words, and the trials containing repeat items were randomized for each participant.

Procedure

Participants viewed two blocks of words, which lasted around 35 minutes each. Both blocks consisted of 1360 trials. Participants were given a short break (5 minutes) between the two blocks. As in experiment 3, stimuli were presented for 2.6 seconds, and phase scrambled images were used as masked separators and presented for 400 ms between each image. The stimuli, approximately 7.5° by 7.5° of visual angle, were presented one at a time in the center of the screen. Participants were instructed to hit the space bar if the current word had previously been presented. They did not need to do anything on trials where they judged the current word to be new. Participants only received feedback on trials in which they responded. On false alarm trials a red square appeared around the word and on trials which were hits, a green square appeared around the word.

3.1.5 Experiment 5

Participants

Forty-one individuals participated and were compensated with \$15 for their time. All participants were members of the Boston University community, and were recruited via the quick job page. Participants gave informed written consent prior to beginning the study, and the protocol was approved by the Boston University Charles River Institutional Review Board. Participants completed the task at different times at one of two 21.5 inch, mid-2011 iMac computers.

Materials

Stimuli were assembled from the same set of images as used in experiments 1-3. Participants viewed a series of 1360 images, 276 of which were randomly chosen to be repeated. Of those 276 repeated images, 30 were chosen to be repeated immediately (lag 1), 28 were repeated at a lag of 2, and so on out to 12 images repeated at lag 512. Repetitions that occurred across blocks were not considered. For each subject, 27 items were repeated a second time at lags 2, 4, and 16 (9 per lag). For items repeated a second time, the lag between its first presentation and first repetition was restricted to 8, 16, and 64. Both the order of the images, and the trials where repeated items occurred were randomized for each participant.

Procedure

Participants viewed two blocks of words, which lasted around 35 minutes each. Both blocks consisted of 680 trials. Participants were given a short break (5 minutes) between the two blocks. As in experiment 3, stimuli were presented for 2.6 seconds, and phase scrambled images were used as masked separators and presented for 400 ms between each image. The stimuli, approximately 7.5° by 7.5° of visual angle, were

presented one at a time in the center of the screen. Different from previous experiments, participants were instructed to hit the left arrow key if the current image was repeated, and the right arrow key if the current image was new. Participants only received feedback on trials in which they responded. On incorrect trials (false alarms and misses) a red square appeared around the image and on correct trials (hits and correct rejections) a green square appeared around the image.

3.1.6 Experiment 6

Participants

Forty-six individuals participated and were compensated with \$15 for their time. All participants were members of the Boston University community, and were recruited via the quick job page. Participants gave informed written consent prior to beginning the study, and the protocol was approved by the Boston University Charles River Institutional Review Board. Participants completed the task at different times at one of two 21.5 inch, mid-2011 iMac computers.

Materials

Stimuli were assembled from the same set of images as used in experiment 5. Participants viewed a series of 1360 images, 276 of which were randomly chosen to be repeated. Of those 276 repeated images, 30 were chosen to be repeated immediately (lag 1), 28 were repeated at a lag of 2, and so on out to 12 images repeated at lag 512. Repetitions that occurred across blocks were not considered. For each subject, 27 items were repeated five times. For items repeated more than once, the lag between its first presentation and first repetition was restricted to 8, 16, and 64. For all subsequent repetitions, the possible lags were 2, 4, and 16 (9 per lag and number of repetitions). Both the order of the images, and the trials where repeated items occurred were randomized for each participant.

Procedure

Participants viewed two blocks of words, which lasted around 35 minutes each. Both blocks consisted of 1360 trials. Participants were given a short break (5 minutes) between the two blocks. Stimuli were presented for 2.6 seconds, and phase scrambled images were used as masked separators and presented for 400 ms between each image. The stimuli, approximately 7.5° by 7.5° of visual angle, were presented one at a time in the center of the screen. As in experiment 5, participants were instructed to hit the left arrow key if the current image was repeated, and the right arrow key if the current image was new. Participants only received feedback on trials in which they responded. On incorrect trials (false alarms and misses) a red square appeared around the image and on correct trials a green square appeared around the image.

3.1.7 Analyses

Analyses were performed using R statistical software. Both linear and logistic regressions were performed as mixed effects regressions, using the NLME package in R (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2021). All analyses were performed on the base 2 log of the lag. Due to the possibility that responses for items repeated immediately may not necessarily involve retrieval from memory, lag 1 repetitions were not considered in any regressions. Lags larger than 128 were also not considered. Within Subject error bars were calculated using the method outlined in Morey (2008).

All data and code used to perform analyses have been made publicly available on Github and can be accessed at <https://github.com/tcnlab/ConRec>.

3.2 Results

To anticipate the results, the data from all six experiments showed evidence that the time to retrieve a memory, as operationalized by the onset time of the RT distribution, changed systematically with the recency of the probe item. When considering probes repeated once, they were faster and more likely to be recognized the more recently an item was last presented. Analysis of response time distributions showed an effect of recency even for the fastest response times. Analysis of multiple repetitions found that correct RT's decreased with the number of previous presentations. Critically, the time to access a previously presented probe only depended on the recency of its most recent presentation. Further, the effect of recency was the same regardless of the number of repetitions. In all figures, experiment 1 corresponds to black circles and solid lines, experiment 2 corresponds to black triangles and dashed lines, experiment 3 corresponds to blue circles and solid lines, experiment 4 corresponds to blue triangles and dashed lines, experiment 5 corresponds to red circles and solid lines, and experiment 6 corresponds to red triangles and dashed lines.

3.2.1 First Repetitions

Items presented more recently were more likely to be recognized

The average false alarm rate was low across experiments. In experiment 1, the false alarm rate was 1.4 percent, the highest hit rate was at lag 1 at 99.6 percent (corresponding to a d' of 4.85) and lowest at lag 128 at 89.7 percent ($d' = 3.46$). In experiment 2, the false alarm rate remained low across participants at 3.8 percent, hit rate was highest for lag 1 at 96.2 percent ($d' = 3.55$), and lowest for lag 128 with an average hit rate of 69.2 percent ($d' = 2.27$). Experiment 3 was similar to experiment 2, the false alarm rate was 3.0 percent, hit rate was the highest for lag 1 at 91.9 percent ($d' = 3.28$), and lowest for lag 128 with an average hit rate of 69.2

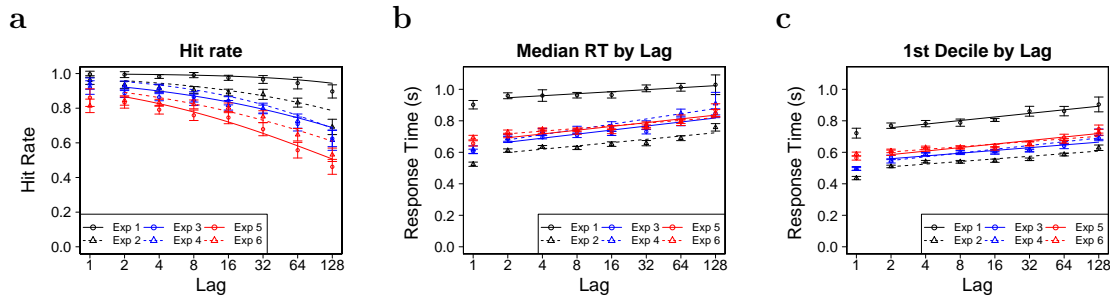


Figure 3-2: The recency effect is robust across experimental conditions. **a.** Hit Rate as a function of \log_2 lag for Experiments 1-6. The hit rate goes down with lag. There is a more pronounced drop in the hit rate at higher lags. **b.** Median response time as a function of \log_2 lag. Median response time increased linearly with the logarithm of lag. **c.** The 1st decile as a function of \log_2 lag. 1st decile response times increased linearly with the logarithm of the lag at a similar rate as median response times. Error bars in all figures represent the 95% confidence interval of the mean across participants normalized using the method described in Morey (2008).

percent ($d' = 3.28$). In experiment 4, which used words as stimuli, the false alarm rate was higher than in previous experiments at 6.9 percent, hit rate was highest for lag 1 at 94.2 percent ($d' = 3.06$), and lowest for lag 128 with an average hit rate of 61.5 percent ($d' = 1.78$). In experiment 5, which used the same stimuli as experiments 1-3 but required a new/old response, the false alarm rate was similar to those experiments at 2.0 percent, but overall hit rate was lower. Hit rate was highest for lag 1 at 86.1 percent ($d' = 3.12$) and lowest for lag 128 with an average hit rate of 46.2 percent ($d' = 1.94$). Experiment 6, which also required new/old responses for each trial was resulted in a similar false alarm rate of 2.6 percent, hit rate was highest for lag 1 at 84.4 percent ($d' = 2.96$), and lowest for lag 128 with an average hit rate of 46.2 percent ($d' = 2.03$). Subjects were able to distinguish old probes from new probes successfully out to at least six minutes.

Average subject hit rate as a function of log lag is plotted in Figure 3-2a. There

Hit Rate	Exp. 1	Exp. 2	Exp. 3
Intercept	6.04 ± 0.47 $z = \mathbf{12.75}, p < 0.001$	3.45 ± 0.18 $z = \mathbf{19.18}, p < 0.001$	2.74 ± 0.16 $z = \mathbf{17.54}, p < 0.001$
Slope	-0.46 ± 0.07 $z = \mathbf{-6.72}, p < 0.001$	-0.31 ± 0.02 $z = \mathbf{-13.88}, p < 0.001$	-0.28 ± 0.02 $z = \mathbf{-15.27}, p < 0.001$
	Exp. 4	Exp. 5	Exp. 6
Intercept	3.39 ± 0.22 $z = \mathbf{15.17}, p < 0.001$	2.17 ± 0.15 $z = \mathbf{14.59}, p < 0.001$	2.38 ± 0.15 $z = \mathbf{16.22}, p < 0.001$
Slope	-0.37 ± 0.02 $z = \mathbf{-17.53}, p < 0.001$	-0.31 ± 0.02 $z = \mathbf{-16.90}, p < 0.001$	-0.28 ± 0.02 $z = \mathbf{-15.85}, p < 0.001$

Table 3.2: Slope and intercept values from a logistic mixed effects regression of lag on hit rate for first repetition items. All tests were performed on the base 2 logarithm of lag. Bold z-scores indicate significance at the $p < 0.05$ level. Across all six experiments, hit rates decreased as a function of lag.

was a recency effect in accuracy for all experiments. In order to quantify the decrease in accuracy for increasing lags, we performed a mixed effects logistic regression on hit rate, treating log lag as a fixed effect and subject as a random effect. Table 3.2 shows the results of this analysis for each experiment. Across all six experiments, hit rate significantly decreased with each doubling of lag ($p < 0.001$). Subjects were less likely to correctly recognize repeated items the further in the past they were presented.

More recent items were recognized faster than older items

Figure 3-2b shows the average median correct response time across subjects as a function of log lag. There was a recency effect on median response times; more recently presented items were recognized faster than less recent items. To confirm the existence of this recency effect, a mixed effects linear regression, treating subject as a random effect and log lag as a fixed effect was performed. Table 3.4 shows the results of this analysis for each experiment. Across all six experiments, median response times significantly increased for each doubling of lag ($p < 0.001$). Subjects recognized items faster the more recently they were presented.

As shown in Figure 3-1, the two classes of models that explain the recency effect make differing predictions about how the leading edge of their distributions should

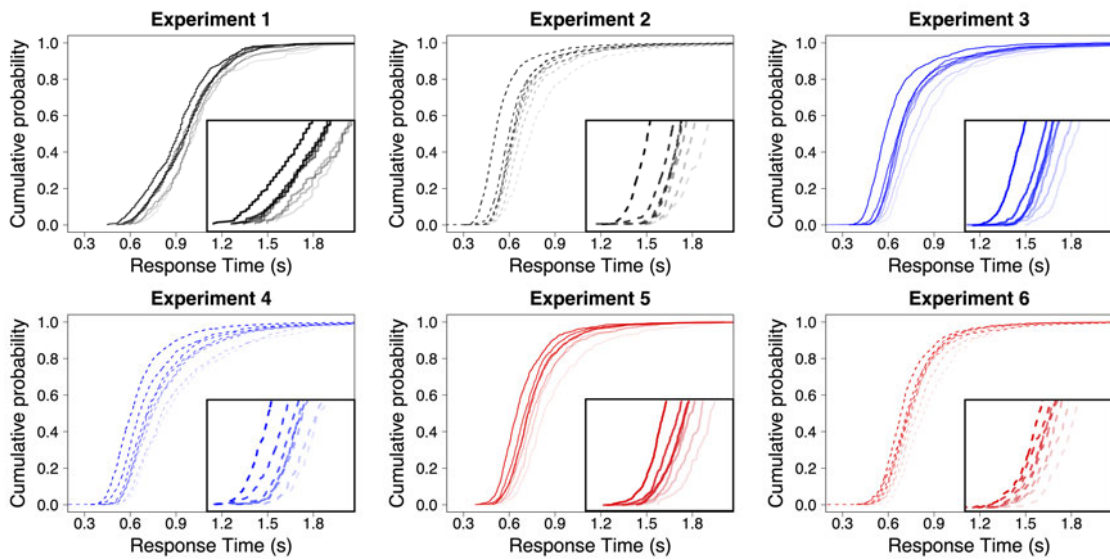


Figure 3-3: Time to access memory changed systematically with lag in all six experiments. Unsmoothed across-participant cumulative response distributions for each lag. Shorter lags correspond to darker lines. Note that the cumulative distributions shift with decreasing recency. The inset consists of a zoomed in view of the rising points of the distributions. For all six experiments, there is consistent evidence that newer items are available before older items.

Median	Exp. 1	Exp. 2	Exp. 3
Intercept	0.934 ± 0.042 $t(83) = \mathbf{22.22}, p < 0.001$	0.578 ± 0.013 $t(221) = \mathbf{45.05}, p < 0.001$	0.639 ± 0.021 $t(227) = \mathbf{30.16}, p < 0.001$
Slope	0.013 ± 0.003 $t(83) = \mathbf{4.54}, p < 0.001$	0.021 ± 0.002 $t(221) = \mathbf{13.65}, p < 0.001$	0.026 ± 0.002 $t(227) = \mathbf{13.20}, p < 0.001$
	Exp. 4	Exp. 5	Exp. 6
Intercept	0.646 ± 0.022 $t(197) = \mathbf{28.88}, p < 0.001$	0.669 ± 0.020 $t(179) = \mathbf{32.85}, p < 0.001$	0.700 ± 0.016 $t(209) = \mathbf{42.55}, p < 0.001$
Slope	0.033 ± 0.003 $t(197) = \mathbf{9.69}, p < 0.001$	0.024 ± 0.002 $t(179) = \mathbf{10.52}, p < 0.001$	0.017 ± 0.001 $t(209) = \mathbf{11.98}, p < 0.001$

Table 3.3: Slope and intercept values from a linear mixed effects regression of lag on median response times for first repetition items. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times increased as a function of lag.

vary with lag. Figure 3-3 shows the empirical cumulative response time distributions for all lags. The curves rise at different points, with smaller lags rising earlier than larger lags, indicating an effect of lag for even the fastest responses. As leading edge of the distribution is well described by the first decile (Ratcliff & Smith, 2004; Smith & Ratcliff, 2009), evidence of a recency effect at this point in the distribution would indicate a shift in the RT distributions. Figure 3-2c shows the average first decile correct response time across subjects as a function of log lag. There was a recency effect present for the fastest responses. A mixed effects linear regression on 1st decile response times, treating subject as a random effect and log lag as a fixed effect confirmed this visual impression. Table 3.4 shows the results of this analysis for each experiment. For each doubling of lag, first decile response times significantly increased. These findings demonstrate that there was a recency effect for even fastest response times, consistent with the shift predicted by a recency dependent retrieval process.

1st Decile	Exp. 1	Exp. 2	Exp. 3
Intercept	0.734 ± 0.033 $t(83) = \mathbf{22.05}, p < 0.001$	0.492 ± 0.008 $t(221) = \mathbf{60.63}, p < 0.001$	0.542 ± 0.012 $t(227) = \mathbf{45.17}, p < 0.001$
Slope	0.023 ± 0.002 $t(83) = \mathbf{9.29}, p < 0.001$	0.017 ± 0.001 $t(221) = \mathbf{15.65}, p < 0.001$	0.018 ± 0.001 $t(227) = \mathbf{14.22}, p < 0.001$
	Exp. 4	Exp. 5	Exp. 6
Intercept	0.524 ± 0.012 $t(197) = \mathbf{41.91}, p < 0.001$	0.563 ± 0.014 $t(179) = \mathbf{39.26}, p < 0.001$	0.586 ± 0.013 $t(209) = \mathbf{44.59}, p < 0.001$
Slope	0.024 ± 0.002 $t(197) = \mathbf{14.76}, p < 0.001$	0.022 ± 0.002 $t(179) = \mathbf{13.66}, p < 0.001$	0.017 ± 0.001 $t(209) = \mathbf{13.23}, p < 0.001$

Table 3.4: Slope and intercept values from a linear mixed effects regression of lag on first decile response times for first repetition items. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times increased as a function of lag.

Non-parametric analyses of response time distributions indicate that more recent memories are available earlier than less recent memories

In order to better formalize this shift, particularly for the fastest responses, we calculated the slope of each subject’s response times as a function of log lag, which we refer to as its “Lag Modulation Factor”, and measured it at different quantiles. As illustrated in Figure 3-1, a strength-based account predicts that Lag Modulation Factor is zero for the fastest responses, and increases later in the distribution. If instead, recency determines when a memory can be compared to the probe, then Lag Modulation Factor is non-zero, even for the fastest responses. The results of the Lag Modulation Factor analysis offered substantial support to this view. For all six experiments, a mixed effects linear regression of lag modulation factor onto quantile, treating subject as a random effect and quantile as a fixed effect was fit to the data. Table 3.5 shows the results of this analysis for the six experiments. Critically, in all six experiments lag modulation factor was significant at the intercept ($p < 0.001$), indicating there was an effect of lag on response time for even the fastest responses. The rate at which lag modulation factor changed throughout the distribution varied across experiments, casting further doubts on the ability of a composite model to account for this pattern of results. In experiments 2 ($p < 0.01$), 3 and 4 ($p < 0.001$), lag

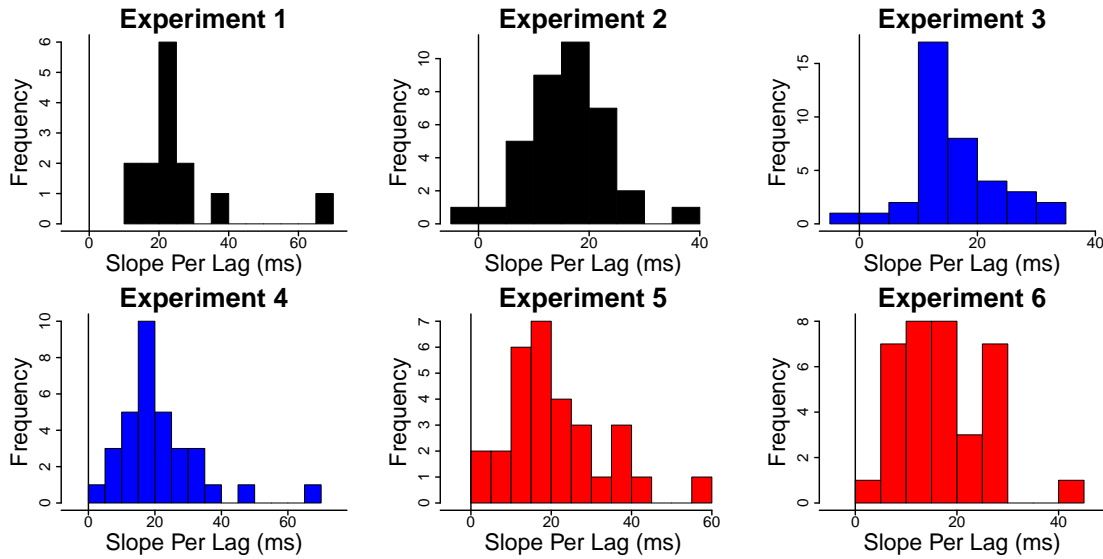


Figure 3-4: Recency impacted response times for even the fastest responses. Histograms of each subject’s slope per lag (ms) for the fastest responses as calculated by the Lag Modulation Factor analysis. Across all six experiments, subjects had significant recency effects at the start of their RT distributions.

modulation factor significantly increased, consistent with a drift rate that is slower for less recent items. This is complicated however by the finding that lag modulation factor did not significantly change with quantile in experiments 5 and 6 ($p > 0.1$), and significantly decreased in experiment 1 ($p < 0.001$). This pattern of results is also striking when looking at the subject level. Figure 3-4 plots the intercept values for each subject’s lag modulation factor as calculated by a linear regression of Lag Modulation Factor onto quantile. With the exception of a single subject in experiments 2 and 3, all participants in all experiments had a positive lag modulation factor at the intercept. Taken together, these results provide strong support that the recency effect in continuous recognition is driven by a change in the time at which items become available.

Lag Modulation Factor	Exp. 1	Exp. 2
Intercept	25.25 ± 3.72 $t(97) = \mathbf{6.78}, p < 0.001$	16.08 ± 1.12 $t(258) = \mathbf{14.41}, p < 0.001$
Slope	-26.69 ± 4.12 $t(97) = \mathbf{-6.47}, p < 0.001$	6.29 ± 1.84 $t(258) = \mathbf{3.41}, p < 0.01$
	Exp. 3	Exp. 4
Intercept	15.92 ± 1.37 $t(265) = \mathbf{11.62}, p < 0.001$	21.44 ± 3.25 $t(230) = \mathbf{6.61}, p < 0.001$
Slope	-26.69 ± 4.12 $t(265) = \mathbf{7.44}, p < 0.001$	24.16 ± 4.27 $t(230) = \mathbf{5.66}, p < 0.001$
	Exp. 5	Exp. 6
Intercept	21.70 ± 2.71 $t(209) = \mathbf{8.02}, p < 0.001$	16.97 ± 1.37 $t(244) = \mathbf{12.38}, p < 0.001$
Slope	3.23 ± 2.37 $t(209) = 1.36, p = 0.17$	-1.16 ± 1.92 $t(244) = -0.60, p = 0.55$

Table 3.5: The slopes and intercepts of the linear mixed effects regression on Lag Modulation Factor as a function of Quantile across the six experiments for first repetition items. Bold t-scores indicate significance at the $p < 0.05$ level. The intercepts calculated from the lag modulation factor analysis were significant in all six experiments.

3.2.2 Multiple Repetitions

Analysis of first repetition responses showed evidence that recency changes response time distributions via a shift in the distribution, consistent with a change in the time to access memory. In order to consider the hypothesis that the recency effect was attributable to changes in encoding, we examined response times for probes repeated more than once.

Only the most recent lag impacted response times for items repeated multiple times

Response times depended only on the most recent lag. We performed a mixed effects linear regression of median response times for items repeated a second time, treating the most recent lag (lag_2) and the lag between the first and second presentations of the image (lag_1) as fixed effects and subject as a random effect (Table 3.6). In experiments 1-5, this analysis found that doubling lag_2 significantly increased median response times, while doubling lag_1 did not have a significant effect. In experiment 6, while

doubling lag_2 resulted in a significant increase in response times, unlike in previous experiments, doubling lag_1 also resulted in a significant increasing in response times of. Subsequent analyses suggest this divergence in results was an artifact of how the lists were assembled. In experiment 6, which was kept consistent with experiments 2-5, lag_1 was confounded with number of repetitions. For items repeated a second time, lag_1 could equal 8, 16, and 64, for all subsequent repetitions, lag_1 could equal 2, 4, and 16. It is possible that response times do not increase with lag_1 , but instead our analysis is capturing that second repetition items are slower than items repeated 3 or more times. In order to determine if this was the case, two follow up analyses were performed. When considering only items repeated three or more times, removing the confound of lag_1 and repetition, only the most recent lag impacted response times. As before, we performed a mixed effects linear regression of median response times for items repeated a three or more times, treating the most recent lag (lag_2) and the lag between the previous repetition and its previous presentation (lag_1) as fixed effects and subject as a random effect. We found a significant effect of lag_2 on RT ($0.011 \pm 0.002, t(278) = 6.77, p < 0.001$), but no significant effect of lag_1 ($-0.001 \pm 0.002, t(488) = -0.31, p = 0.75$). As an additional control, we performed the same mixed effects linear regression for median response times on items repeated twice or more as before, but included a fixed variable that was 1 if a trial contained the second repetition of an item, and zero for all other repetitions in order to separate the effect of (lag_1) from a repetition effect. We found significant effects of lag_2 ($0.008 \pm 0.002, t(592) = 4.25, p < 0.001$) and repetition ($0.058 \pm 0.006, t(592) = 9.30, p < 0.001$), but lag_1 did not significantly effect response times ($-0.002 \pm 0.002, t(592) = -1.36, p = 0.17$). These subsequent analyses suggest that our results in experiment 6 are in line with the previous experiments, and indicate that the recency effect is not the result of improved encoding for more recently presented items.

Median	Exp. 1	Exp. 2	Exp. 3
Intercept	0.871 ± 0.041 $t(393) = \mathbf{21.11}, p < 0.001$	0.534 ± 0.015 $t(294) = \mathbf{35.71}, p < 0.001$	0.605 ± 0.022 $t(302) = \mathbf{27.04}, p < 0.001$
Lag 2	0.021 ± 0.005 $t(393) = \mathbf{4.34}, p < 0.001$	0.012 ± 0.002 $t(294) = \mathbf{5.37}, p < 0.001$	0.012 ± 0.003 $t(302) = \mathbf{4.43}, p < 0.001$
Lag 1	-0.004 ± 0.004 $t(393) = -0.83, p = 0.41$	-0.001 ± 0.002 $t(294) = -0.54, p = 0.59$	-0.004 ± 0.003 $t(302) = -1.36, p = 0.17$
	Exp. 4	Exp. 5	Exp. 6
Intercept	0.634 ± 0.027 $t(261) = \mathbf{23.06}, p < 0.001$	0.649 ± 0.025 $t(235) = \mathbf{26.30}, p < 0.001$	0.573 ± 0.017 $t(488) = \mathbf{34.30}, p < 0.001$
Lag 2	0.010 ± 0.005 $t(261) = \mathbf{2.00}, p = 0.046$	0.012 ± 0.004 $t(235) = \mathbf{3.26}, p = 0.001$	0.007 ± 0.002 $t(488) = \mathbf{4.83}, p < 0.001$
Lag 1	0.001 ± 0.005 $t(261) = 0.30, p = 0.78$	-0.005 ± 0.004 $t(235) = -1.48, p = 0.14$	0.007 ± 0.002 $t(488) = \mathbf{3.56}, p < 0.001$

Table 3.6: Slope and intercept values from a linear mixed effects regression of the lag between an item’s second and third presentation (lag_2) and the lag between an item’s first and second presentation (lag_1) and on median response times for items repeated a second time. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times increased as a function of lag_2 but did not systematically vary for lag_1 .

Items repeated multiple times are retrieved faster, but the effect of recency is consistent across repetitions

Figure 3-5 shows median subject response times as a function of log lag and number of repetitions. Subjects were faster for items repeated more than once, but the effect of recency was consistent across repetitions. A linear mixed effects regression on median response times was performed treating subject as a random effect and log lag, number of repetitions, and the interaction of lag and repetition as fixed effects (Table 3.7). The same pattern of results emerged for all six experiments, there were significant main effects of log lag and repetition, but their interaction was not significant. That is, while participants were faster to respond the more times an item was presented, and were slower for less recent items, the delay to access those less recent items remained consistent regardless of how many times an item was presented. Taken together, these results indicate that the recency effect does not appear to be the result of better fidelity or improved encoding for more recent items, but rather arises

Median	Exp. 1	Exp. 2	Exp. 3
Intercept	0.934 ± 0.042 $t(179) = \mathbf{22.44}, p < 0.001$	0.602 ± 0.012 $t(182) = \mathbf{50.35}, p < 0.001$	0.674 ± 0.020 $t(187) = \mathbf{33.34}, p < 0.001$
Slope	0.013 ± 0.004 $t(179) = \mathbf{3.47}, p < 0.001$	0.013 ± 0.002 $t(182) = \mathbf{5.48}, p < 0.001$	0.012 ± 0.003 $t(187) = \mathbf{4.56}, p < 0.001$
Repetition	-0.069 ± 0.023 $t(179) = \mathbf{-2.96}, p = 0.004$	-0.067 ± 0.009 $t(182) = \mathbf{-7.52}, p < 0.001$	-0.080 ± 0.009 $t(187) = \mathbf{-8.52}, p < 0.001$
Rep X Slope	0.004 ± 0.005 $t(179) = 0.84, p = 0.40$	-0.004 ± 0.003 $t(182) = -1.13, p = 0.26$	0.002 ± 0.004 $t(187) = 0.52, p = 0.60$
	Exp. 4	Exp. 5	Exp. 6
Intercept	0.670 ± 0.020 $t(162) = \mathbf{34.06}, p < 0.001$	0.671 ± 0.016 $t(147) = \mathbf{42.30}, p < 0.001$	0.686 ± 0.016 $t(487) = \mathbf{42.45}, p < 0.001$
Slope	0.024 ± 0.004 $t(162) = \mathbf{5.87}, p < 0.001$	0.007 ± 0.002 $t(147) = \mathbf{3.03}, p = 0.003$	0.009 ± 0.003 $t(487) = \mathbf{3.36}, p < 0.001$
Repetition	-0.044 ± 0.016 $t(162) = \mathbf{-2.82}, p = 0.006$	-0.048 ± 0.009 $t(147) = \mathbf{-5.33}, p < 0.001$	-0.039 ± 0.003 $t(487) = \mathbf{-13.89}, p < 0.001$
Rep X Slope	-0.007 ± 0.006 $t(263) = -1.15, p = 0.25$	0.004 ± 0.003 $t(147) = 1.15, p = 0.25$	0.001 ± 0.001 $t(487) = 0.88, p = 0.38$

Table 3.7: Slope and intercept values from a linear mixed effects regression of lag and number of repetitions on median response times. All tests were performed on the base 2 logarithm of lag. Bold t-scores indicate significance at the $p < 0.05$ level. Across all six experiments, response times decreased with number of repetitions, but the effect of lag was the same regardless of number of repetitions.

at the comparison stage.

The time that items repeated multiple times become available depends on its recency

As demonstrated via Lag Modulation Factor, for items repeated once, the time an item becomes available depends on its recency (Figure 3.3). Although experiments 1-5 do not have enough trials to calculate a reliable Lag Modulation Factor for items repeated multiple times (9 trials per lag and subject), we can perform this on the data from experiment 6 (36 trials per lag and subject). We performed a mixed effect linear regression of Lag Modulation Factor onto quantile, treating subject as a random effect and quantile as a fixed effect. We found that Lag Modulation Factor was significantly different from zero at the intercept, $9.82 \pm 1.11, t(244) = 8.87, p < 0.001$, but did not significantly change as a function of quantile $1.29 \pm 1.85, t(244) = 0.70, p = 0.49$. That is, while there was an effect of recency for items repeated multiple times, for even the

fastest responses, the effect of recency did not change throughout the distribution of responses. Taken together, these results offer strong support for the hypothesis that the recency effect is caused by the recovery of a temporal context, where the time to retrieve a context depends on its recency.

3.3 General Discussion

It has long been known that response times in a continuous recognition experiment increase with the lag to the probe. If memory search requires accessing a timeline to find the appropriate memory, then the RT increase could be associated with a shift in the distribution. These six experiments showed that lag consistently shifted RT distributions. Consistent with this account, RT's to second repetitions depended only on the most recent lag, as if the search terminated upon finding the first match. Although RTs were faster to second repetitions overall, the effect of the most recent lag on second repetition RTs was the same as the effect of lag on first presentation RTs (Figure 3.5). Consistent with earlier findings (e.g., Hockley, 1982), our results showed a sublinear shift with lag. This paper's results are roughly consistent with a logarithmic increase in RT as a function of lag; each doubling of lag resulted in a shift of approximately 16-26 ms in the RT distribution.

At first glance, the results of this study are consistent with sequentially accessing a logarithmically-compressed timeline. There is extensive evidence for self-terminating serial search models in short-term memory tasks (Hacker, 1980; Hockley, 1984; McElree & Doshier, 1993; Sternberg, 2016). There is also evidence consistent with scanning along a timeline in judgement of recency tasks (D. L. Hintzman, 2010; Tiganj, Singh, Esfahani, & Howard, 2021). To be clear however, we are not suggesting that what we see here is scanning. The effect of doubling lag in our experiments resulted in an increase in RT of 16-25 ms, much faster than the increases seen in tasks believed to

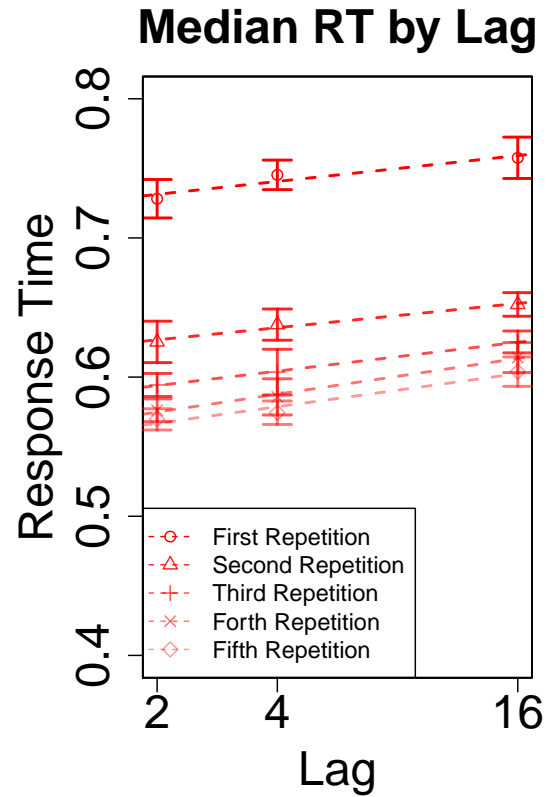


Figure 3-5: The effect of recency on response times was the same regardless of repetitions. Median response time as a function of the most recent lag and number of repetitions. To the extent the lines are parallel, it means that the effect of recency on median RT's was the same for across repetitions. Analyses reported in the text demonstrated that only the most recent lag affected the RT of old probes repeated twice.

show scanning.

Most previous item recognition studies however have found evidence for parallel memory access, not sequential, in study-test recognition (Nosofsky, Little, Donkin, & Fific, 2011; Ratcliff & Murdock, 1976; McElree & Doshier, 1989; Hockley, 1984; Nosofsky, Cox, Cao, & Shiffrin, 2014). Several potentially important methodological differences may account for the discrepancy between those studies and this paper's results. This experiment used continuous recognition rather than the study-test procedure (Nosofsky et al., 2011; Ratcliff & Murdock, 1976; McElree & Doshier, 1989; Hockley, 1984; Nosofsky et al., 2014). In addition, more recent lags were tested more frequently than more remote lags, new probes occurred far more often than repeated probes, and repeated probes could not appear in adjacent trials. The question of which combination of these methodological differences accounts for evidence supporting a timeline is a significant one that merits further investigation. It is worth noting that there is no reason in principle that a compressed timeline could not be accessed in parallel (Howard et al., 2015), whereas it is not clear how (or why) a composite representation could allow for changes in the time at which items become available. Notably, if a logarithmically-compressed timeline is accessed in parallel, the recency effect for the strength of match would fall off like a power law (Howard et al., 2015), much like the change Donkin and Nosofsky (2012) observed experimentally in drift rate.

Logarithmic compression is ubiquitous in psychology. Behaviorally, one way this manifests is via the Weber-Fechner law, the finding that the perceived intensity of stimulus changes with the log of its actual intensity (Fechner, 1860/1912). It has also been well established that compression is seen in neural populations. For example, the representation of visual space is logarithmically compressed as a function of distance from the fovea (E. L. Schwartz, 1980; Van Essen, Newsome, & Maunsell, 1984).

Compression is also seen in time cells, neurons with temporal receptive fields which consistently fire at a specific time during a delay. Time cells show compression in that more cells have receptive fields earlier in the delay, and later firing cells have larger receptive fields (Tiganj et al., 2018; Cruzado et al., 2020). A recent analysis of time cells in rodent hippocampus, an area critical to episodic memory function, demonstrated that time cell compression is in fact logarithmic (Cao et al., 2021), such that, their distance in neural space goes up with log lag.

Episodic recollection is believed to be the result of occurs following the recovering a previous spatiotemporal context. This recovery results in behavioral changes, such as the contiguity effect. Previous recognition studies have found evidence that following episodic recollection, there is a boost in performance for items presented close in time to the just recovered item (G. Schwartz et al., 2005; Folkerts et al., 2018). In addition to behavioral effects, episodic recollection is accompanied by a neural contiguity effect as well (Folkerts et al., 2018), such that the similarity of the population at test is most similar to when the just recollected item was first presented, but also shows a high similarity to other items presented around the same time at study. The present datasets do not allow for us to measure a contiguity effect, as adjacent trials did not contain repetitions and we do not have a measure to separate “episodic recollection” trials from more familiarity base recognition. Despite this, our results are consistent with the hypothesis that the time to recover a temporal context depends on its recency in log space.

Chapter 4

Rapid presentation rate negatively impacts the contiguity effect in free recall.

4.1 Introduction

Cognitive neuroscientists have hypothesized that the successful retrieval of an episodic memory is accompanied by a “jump back in time,” a recovery of the previous memory’s spatiotemporal context (Tulving, 1983). In free recall studies, this recovery manifests as the contiguity effect, wherein following the successful recall of an item, the next item to be recalled is more likely to be a close temporal neighbor than a more distant one (Kahana, 1996). This distance is measured as lag, a directed distance between items in a study list. For example, in the list “absence, hollow, pupil, river, darling”, the lag from absence to river is +3, while the lag from darling to pupil is -2 . In free recall studies the contiguity effect is typically asymmetric, such that forward transitions are more likely to take place than backward transitions of the same distance. This effect is robust, appearing across a variety of methodological manipulations (Kahana, 2012; Healey & Kahana, 2014). For instance, the contiguity effect is observed with more or less the same properties for lists of different modalities (Kahana, 1996), when rehearsal is discouraged (Howard & Kahana, 1999), and when words are widely separated in time (Howard et al., 2008; Unsworth, 2008). Healey and Kahana (2014) noted that the contiguity effect was observed for every individual participant in a free recall study of 126 subjects. Thus far, dramatic effects on the

contiguity effect in free recall have primarily been observed comparing patient populations; older adults and memory disordered individuals show impaired contiguity effects (Kahana, Howard, Zaromb, & Wingfield, 2002; Palombo, Di Lascio, Howard, & Verfaellie, 2019).

Beyond the contiguity effect, free recall contains many other well-explored patterns of behavior. Individuals exhibit a strong recency effect during immediate free recall tests (Glanzer & Cunitz, 1966). In addition, participants exhibit a primacy effect such that items at the beginning of a studied list are more likely to be recalled (B. B. Murdock, 1962). Both primacy and recency effects are observed in the initiation of free recall, and are both also robustly observed in the probability of first recall, a measure of the serial position curve considering only the first recall (Hogan, 1975; Laming, 1999). The relative strength of primacy and recency is not constant however (B. B. Murdock, 1962). For example, Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, and Usher (2005) found that presentation rates affect the relative strength of primacy and recency, with primacy becoming more prevalent as the presentation rate is increased.

The ubiquity of the contiguity effect in free recall presents something of a challenge for models of memory encoding—if nothing affects the contiguity effect, it makes it more difficult to understand how it comes about. Conversely, if we knew boundary conditions on the contiguity effect it would perhaps shed light on the processes supporting the binding of experiences presented close together in time. In this study we explore the effects of increasing presentation rates on the contiguity effect. If the contiguity effect is disrupted at a particular rate, that suggests the time scale over which the encoding processes necessary for temporal binding take place.

Considerations from the ERP literature and rapid serial visual presentation (RSVP) literature inform the time scale over which contiguity might be disrupted. A to-be-

remembered stimulus typically evokes a P300 waveform approximately 500 ms in duration that is thought to represent the updating of memory representations, even when the stimulus duration itself is on the order of 2 seconds (Donchin, 1981). At presentation rates approaching 10 Hz, there is evidence that individual list items are no longer processed as discrete items, and instead are merged into a single extended cognitive event. For example, individual items in 10 Hz lists receive very low hit rates in an immediate recognition test even when the stimuli are never-before-seen natural images (Potter & Levy, 1969). This poor performance is in stark contrast to the excellent recognition memory for long series of images at slower rates of presentation (Standing, 1973; Brady et al., 2008). However despite this lack of memorability, it is also clear that each item in a 10 Hz stream is processed to some degree, since it is possible to detect specific target items with high probability (Potter, 1976). If the processing of individual items in a list undergoes a qualitative change as the presentation rate is increased to the point at which the representations blend together, then the CRP, primacy effect, and recency effect may be altered. For example, the CRP effect may depend on the ability to place individual items into a discrete temporal representation, and thus it may disappear with faster presentation rates. The probability of first recall could also be altered, since a long-running stream of rapidly presented items imposes a sequential cost on subsequent items due to encoding interference from previous items (Wyble, Bowman, & Nieuwenstein, 2009).

4.2 Methods

4.2.1 Participants

Three hundred and thirty undergraduates from Syracuse University participated in this study. Participants were excluded if they failed to recall a correct word in at least one trial ($n = 15$), and if they did not perform all three conditions ($n = 7$).

Data from 308 participants were used in subsequent analyses.

4.2.2 Procedure

Participants took part in 18 trials. Each trial consisted of 20 words from the Toronto Noun Pool (Friendly et al., 1982). Words were visually displayed at three presentation rates: 2 Hz, 4 Hz, and 8 Hz. Participants completed six trials in each condition. Trial order was randomized. Before the start of a trial, participants viewed a bar that discretely rotated at the same rate that words would be presented to help orient them to the upcoming trial (e.g., before a 2 Hz trial the bar would move twice every second). Following the presentation of the list, participants were prompted to verbally recall as many words as possible from the list. Responses were recorded and later parsed using a semi-automatic speech parsing algorithm.

4.2.3 Analysis

We first examined whether presentation rate affected the average number of valid recalls in a trial. This was done with a repeated measures ANOVA. Post-hoc paired permutations (5000 iterations) and Cohen's D effect sizes on mean recalls were then performed to determine significant differences. Serial position curves (SPC) were computed to show the overall probability of a word being recalled based on its position in the list for each participant. We examined whether the recency and primacy effects changed as a function of presentation rate. We performed a paired permutation test in order to predict the difference in the probability of recall for the first and last items in the list (i.e., probability for position 20 minus probability for position 1).

The probability of first recall (PFR) was calculated by dividing the number of times each serial position was recalled first by the total number of first recalls. We then averaged these probabilities across participants per condition. Finally, we calculated the conditional response probability (CRP) for each lag by dividing the number

of correct recall transitions at a given lag by the total number of possible correct transitions at that lag. In order to control for serial position effects, which differed across conditions, we restricted the lag-CRP analysis to transitions within the middle of the list where probability of recall was approximately equal across presentation rates. In order to test for differences in the CRP at each lag across conditions, we performed a number of mixed-effects logistic regressions. We estimated the CRP as a function of the interaction between the following fixed-effects predictors: absolute lag, its direction (backwards or forwards from the previously recalled item), and presentation rate. We report Z- and T-scored coefficients for all mixed-effects models.

4.3 Results

To summarize the results, memory performance was reduced at faster presentation rates. We replicated previous findings with respect to changes in the serial position curve at fast presentation rates. Critically, the contiguity effect, even measured at serial positions that avoided contributions from primacy and recency, was severely disrupted at fast presentation rates.

4.3.1 As Presentation Rate Increases, Fewer Words are Recalled

As shown in Figure 4-1, the total number of words recalled decreased as presentation rates increased (2 Hz: mean = 3.54, SD = 0.85; 4 Hz: mean = 2.86, SD = 0.662; 8 Hz: mean = 2.41, SD = 0.62; ANOVA: $F(2, 614) = 309.2, p < 0.001$). Post-hoc paired permutations confirmed these results, showing that the presentation rate of 2 Hz yielded significantly higher number of recalls than 4 Hz ($p < 0.001$, Cohen's D = 0.9) and 8 Hz ($p < 0.001$, Cohen's D = 1.5), and that 4 Hz produced significantly more recalls than 8 Hz ($p < 0.001$, Cohen's D = 0.69). This result is consistent with previous findings that faster presentation rates decrease the number of words recalled in a free recall task (B. B. Murdock Jr, 1960).

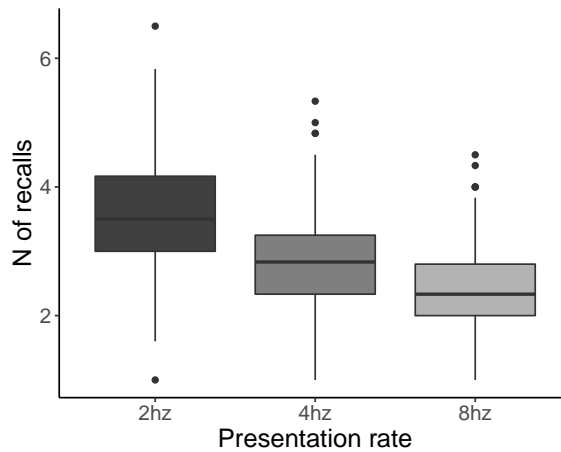


Figure 4-1: A boxplot of median number of words recalled per trial across participants and presentation rates, with interquartile range, 95% confidence intervals and outliers. Participants recalled fewer words as the presentation rate increased.

4.3.2 Increasing Presentation Rates Increases the Primacy Effect and Decreases the Recency Effect

Participants were more likely to begin recall by reporting a word at the beginning or end of the list (Figure 4-2). As the presentation rate increased, participants initiated recall less frequently at the end of the list and more frequently at the beginning of the list. This was confirmed by paired permutation tests which indicated that the probability of beginning a recall with the first item in a studied list was greater at 8 Hz than both 4 Hz ($p < 0.001$, Cohen's $D = 0.24$) and 2 Hz ($p < 0.001$, Cohen's $D = 0.43$), and greater for 4 Hz than 2 Hz ($p < 0.001$, Cohen's $D = 0.18$). Conversely, the probability of first recalling the last item in a list was greater for 2 Hz than both 4 Hz ($p = 0.002$, Cohen's $D = 0.16$) and 8 Hz ($p < 0.001$, Cohen's $D = 0.40$), and higher for 4 Hz than 8 Hz ($p < 0.001$, Cohen's $D = 0.25$).

As shown in Figure 4-3, participants showed a higher rate of recalling words from the beginning and end of a list compared to words in the middle (Figure 4-3). Consistent with previous findings, increasing presentation rates resulted in lower recall for

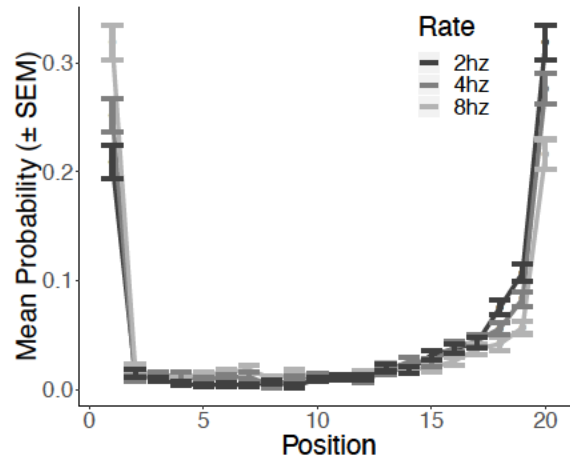


Figure 4-2: Probability of first recall. Participants tended to begin recall by naming an item from the beginning or end of the list. As presentation rates increased, the probability of initiating recall at the end of the list decreased, and the probability of initiating recall with at the beginning of the list increased.

the final item in the list. This was as confirmed by a paired permutation test which found that the probability of recalling the last item in a list was greater for 2 Hz than both 4 Hz ($p < 0.001$, Cohen's $D = 0.50$) and 8 Hz ($p < 0.001$, Cohen's $D = 0.65$), and greater for 4 Hz than 8 Hz ($p = 0.01$, Cohen's $D = 0.15$). In contrast to the PFR (Figure 4-2), increasing presentation rates did not improve the overall probability of recalling the first item in a list. Rather, it was found that the probability of recalling the first item of a list was greater at 2 Hz than 4 Hz ($p = 0.008$, Cohen's $D = 0.17$), and otherwise there were no significant difference (all $p > 0.05$).

4.3.3 The Contiguity Effect Flattens At Higher Presentation Rates

Figure 4-4 shows the number of transitions between each serial position in each of the three conditions. The primacy and recency effects can be readily distinguished, as is the tendency to make remote transitions to the beginning of the list. The contiguity effect can be seen as a slightly darker shade along the diagonal; the forward asymmetry

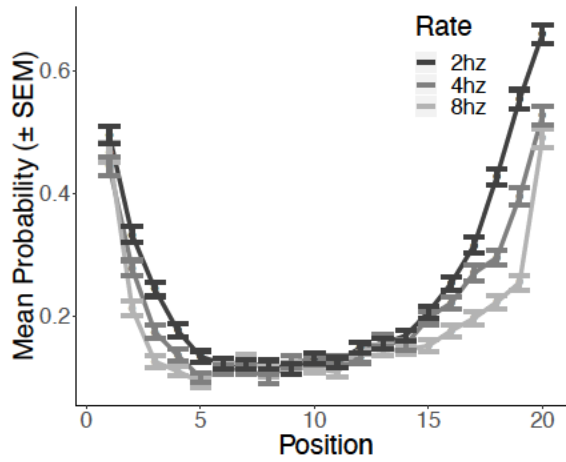
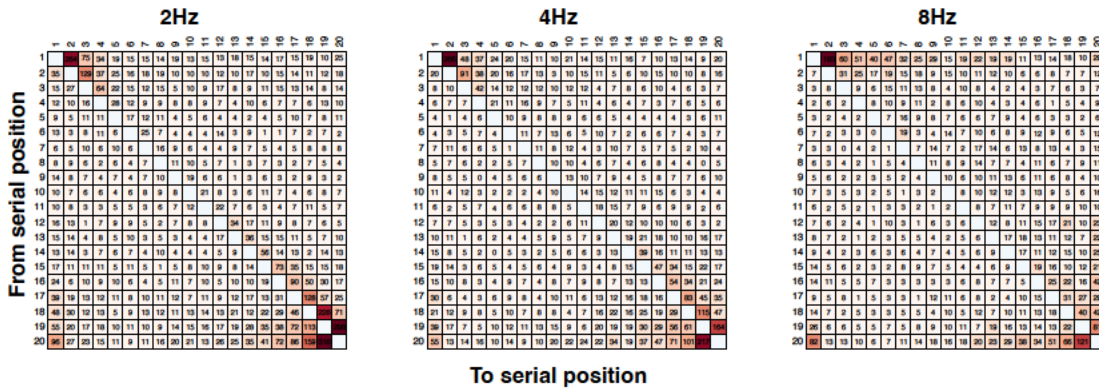


Figure 4.3: Probability of recall as a function of position in study list. Participants showed the highest level of performance for items at the beginning and end of the list. As presentation rate increased, participants showed a tendency to have a lower recency effect in comparison to the primacy effect.



appears as a darker shade just above the diagonal. As expected, the contiguity effect appeared to decrease as presentation rate increased. Because primacy and recency effects are a confound in identifying the contiguity effect we calculated the lag-CRP using only transitions that came from items from the middle of the list (serial positions 7-13).

Figure 4-5 displays the average probability of transitioning from a recalled word to a word at a given lag (with lag 0 corresponding to the diagonal of the matrices in Figure 4), and appears to show a reduction in the temporal contiguity effect as the presentation rate increases. We performed a mixed effect logistic regression to estimate the probability of recall based on absolute lag for each presentation rate separately. This showed that distance from the previously-recalled item significantly decreased the probability of recall for 2 Hz ($z = -9.74, p < 0.001$) and 4 Hz ($z = -4.93, p < 0.001$), but not for 8 Hz ($z = -0.70, p = 0.48$). We then computed another mixed effects logistic regression to test the interaction between absolute lag, its direction (backwards or forwards from the previously recalled item), and the presentation rate. This analysis showed that transitions in the forward direction were more probable than backward transitions ($z = 4.18, p < 0.001$); transitions at more distant lags were less probable ($z = -4.83, p < 0.001$); probabilities were higher for 2 Hz compared to 4 Hz ($z = -2.28, p = 0.02$) and 8 Hz ($z = -3.63, p < 0.001$); the effect of absolute lag was stronger for forward transitions than backwards transitions ($z = -2.71, p < 0.01$), and the effect of lag was stronger for 2 Hz compared to 4 Hz ($z = 2.11, p = 0.03$) and 8 Hz ($z = 3.12, p < 0.01$). All other interactions showed no significant effects (all $p > 0.05$). These results show that increasing the presentation rate of studied words decreases the contiguity effect.

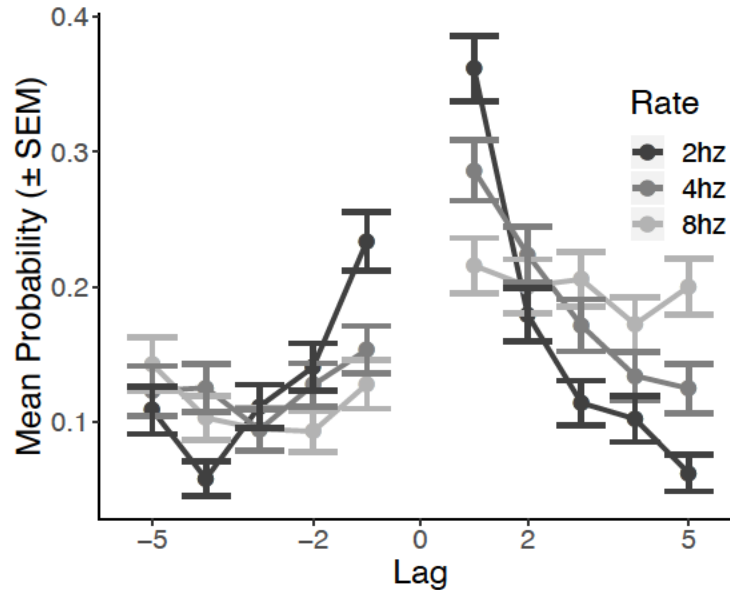


Figure 4.5: Lag-CRP for transitions restricted to serial positions 7-13 to avoid confounds with primacy and recency effects. As presentation rates increase, the contiguity effect weakens but remains asymmetric. At 8 Hz there is no positive evidence for a contiguity effect.

4.4 Discussion

Remembering past events is associated with a jump back in time, manifesting in a higher probability for temporally contiguous elements to be subsequently recalled. In this study, we investigated whether higher presentation rates would negatively impact the temporal contiguity effect. Many of our results were consistent with previous free recall studies. For instance, the average number of words recalled per list decreased as the presentation rate increased. Also, as the presentation rate increased, the recency effect was diminished. While the primacy effect increased in looking at the probability of first recall, there was not a clear effect on the overall probability of recalling the first item. The novel contribution of this paper is the finding that the temporal contiguity effect was disrupted by fast presentation rates, most notably in the 8 Hz condition. These findings suggest that encoding processes taking place on the order

of 125 to 250 ms are important for binding items to their temporal context.

Our results pose questions about the relation of presentation rate and neural coding. Medial temporal lobe theta (3-8 Hz) is related to successful encoding in free recall, particularly when binding elements temporally (Nyhus & Curran, 2010; Sederberg, Kahana, Howard, Donner, & Madsen, 2003). In addition, Guderian, Schott, Richardson-Klavehn, and Düzel (2009) have shown that prediction of successfully-recalled items relies on theta frequency. While presentation rates of 2 Hz and 4 Hz are mostly contained within this frequency band, 8 Hz lies at the upper bound of human theta. It is possible that presenting eight words per second outpaces encoding processes that depend on theta (Hasselmo, Bodelón, & Wyble, 2002), thus explaining why lag-CRPs become weaker for this presentation speed. Examination of encoding and retrieval periods using EEG and ECoG could help address this issue in the future.

Chapter 5

Conclusions

5.1 Summary of dissertation

This dissertation presents neural and behavioral data testing the hypothesis that temporal context has a scale-invariant representation. Chapter 2 presents neural recordings from the monkey entorhinal cortex that demonstrate a population of temporal context cells, a key prediction of this scale-invariant representation. Chapter 3 analyzes response time distributions to demonstrate that the time memories become available during search depends on their recency. Finally, chapter 4 analyzes free recall data showing that rapid presentation rates disrupt the contiguity effect. These results support the assertion that temporal context is scale-invariant and places further constraints on how this context is formed and accessed.

5.1.1 Chapter 2 results

Chapter 2 analyzes neural recordings from the monkey entorhinal cortex as monkeys freely viewed complex natural images for 5 s. We used a maximum likelihood approach to identify time cells (Salz et al., 2016) to find a temporal context cell population. These temporal context cells responded shortly after an image appeared on the screen and relaxed their firing at various rates. Critically, the distribution of relaxation times in this population was non-uniform. More cells had short relaxation times than long relaxation times, a necessary but not sufficient requirement for scale invariance. As expected from a representation of temporal context, the cells contained temporal

information. Temporal context cells also showed limited sensitivity to the image presented on the screen.

5.1.2 Chapter 3 results

Chapter 3 presents the results of a series of six continuous recognition experiments. In all experiments, response times increased with $\log_2(\text{lag})$ out to at least 8 minutes, offering further support that memory is scale-invariant. We did not see a time scale at which the pattern of behavior categorically changed. We analyzed response time distributions and determined that a change in the leading edge of the response time distributions carries the recency effect. The time that memories become available depends on their recency. Items experienced recently become available before less recent items. Each doubling of lag increased this delay in availability by $\tilde{20}$ ms. This effect appeared across a wide range of experimental manipulations. Through analyzing responses to items presented multiple times, we ruled out the possibility that the faster encoding of probes facilitates the recency effect. Even for items presented 6 times, the delay in retrieving less recent items remained.

5.1.3 Chapter 4 results

Chapter 4 presents the results of a free recall task in which words were presented at three possible presentation rates (2, 4, or 8 Hz.). Analyses of the lag-CRP demonstrated that the contiguity effect was intact for lists presented at 2 or 4 Hz, but was disrupted in lists presented at 8 Hz. In the 8 Hz condition, participants were not more likely to recall a word presented nearby after recalling a word. Notably, the asymmetry in the conditional response probabilities remained. In addition, forward transitions were more likely than backward transitions of the same distance. These results show a lower bound on the time scale at which the contiguity effect occurs.

5.2 Future directions

5.2.1 Temporal context cells for the past and future

There is widespread evidence from neural recordings of units that monotonically change their firing in response to an event (Leon & Shadlen, 2003; Lebedev et al., 2008; Kim et al., 2013; Tsao et al., 2018; Ning et al., 2022). This dissertation presents evidence of temporal context cells with a spectrum of decay rates. This result, in conjunction with evidence from other studies (Tsao et al., 2018; Rossi-Pool et al., 2019), offers promising evidence for temporal context cells. Tsao et al. (2018), analyzed single units in rodent LEC while rodents freely explored a series of black and white boxes. Tsao et al. (2018) reported a population of units in LEC that were responsive to different events (e.g., the beginning of a session, entering a white box, entering a black box) and exponentially relaxed their firing. Critically, they reported a wide range of time constants ranging from minutes to hours. As would be required for scale invariance, there were more cells with short time constants than long time constants. Moving away from the MTL, Rossi-Pool et al. (2019) reported a stimulus-specific temporal signal in the monkey dorsal premotor cortex during a task requiring monkeys to remember a vibrotactile stimulus. Although Rossi-Pool et al. (2019) did not quantify decay rates, examination of the raw data indicates that these units have a broad range of relaxation times (Howard, 2022). One limitation of these studies is that none quantify whether the distribution of time constants/relaxation times is scale-invariant. Bayesian methods developed to show that the distribution of time cell peaks is indeed logarithmically distributed (Cao et al., 2021) could be applied to the relaxation times of temporal context cells. To the extent the distribution of time constants is fit by a power law, this would stand as strong evidence that temporal context cells are scale-invariant. There is limited support for this hypothesis in rodent cerebellar slice work, as Guo et al. (2021) reported a spectrum of time constants

consistent with a logarithmic distribution out to 10 s.

Additionally, recent theoretical work has proposed a population of neurons analogous to temporal context cells, but for the future (Howard, Esfahani, Le, & Sederberg, 2023). Rather than all responding shortly after an event and exponentially decaying at a spectrum of rates, these “future temporal context” cells ramp their activity up as a future event approaches such that they peak when the event occurs. Much like evidence from cells that monotonically decay offered limited evidence of temporal context cells, reports of ramping cells support these “Future temporal context cells”. In addition to decaying cells, Tsao et al. (2018) reported cells in rodent LEC that ramped up over time. Like the distribution of time constants for decay cells, there was a spectrum of time constants for ramp cells, spanning minutes to hours. Ramping cells have been reported in an interval reproduction task in rat hippocampus (Hampson & Deadwyler, 2003; Ning et al., 2022), rat PFC (Ning et al., 2022), gerbil PFC (Henke et al., 2021) and mouse anterior lateral motor cortex in a delayed-response task that required planned motor actions (Inagaki, Inagaki, Romani, & Svoboda, 2018). Preliminary secondary analyses of Henke et al. (2021) and Inagaki et al. (2018) suggest that these ramping cells do, in fact, have a range of time constants.

5.2.2 Connections between episodic memory and sharp-wave ripples

Despite widespread agreement that retrieving an episodic memory requires recovering spatiotemporal context, how the brain accomplishes this jump back in time remains a mystery. Sharp-wave ripples (SPW-R), believed to be related to episodic memory (Buzsáki, 2015; Chen et al., 2021), stand as a potential mechanism. SPW-R’s consist of two components; a large depolarization of stratum radiatum in CA1 (the sharp wave) and a high frequency “ripple” oscillation between 110-200 Hz in the pyramidal cell layer (O’Keefe, 1976; Buzsáki, 1996). During ripples, there is a marked increase in firing (Buzsáki, 1986); this phenomenon is frequently referred to as a “population

burst event” (PBE). Ripples have been reported in every mammal that someone has looked for ripples in (Buzsáki, Logothetis, & Singer, 2013), including humans (Bragin, Engel Jr, Wilson, Fried, & Buzsáki, 1999; Le Van Quyen et al., 2010). However, evidence for SPW-R’s in non-mammalians is comparatively scarce (Buzsáki, 2015), but see Vargas, Thorsteinsson, and Karlsson (2012) for evidence of ripples in zebrafish. In contrast to hippocampal theta, which is strongest in rodent CA1 during periods of activity (Vanderwolf, 1969), SPW-R’s occur during periods of immobility, such as when the animal eats, drinks, or grooms (Buzsáki, Leung, & Vanderwolf, 1983).

In humans performing episodic memory tasks, ripple rates increase during encoding and just prior to successful retrieval (Y. Norman et al., 2019; Sakon & Kahana, 2022). Critically, during ripples associated with successful retrieval, iEEG recordings indicate that the brain reinstates the pattern of activity present at encoding (Y. Norman et al., 2019; Vaz, Inati, Brunel, & Zaghoul, 2019). This recovery of a prior brain state suggests that ripples facilitate or at least indicate jumping back in time. Coupled with the finding that older contexts take longer to retrieve in continuous recognition, this hypothesis would suggest that ripples for recent items should occur earlier than less recent items. As far as the author knows, a connection between recency and ripple timing has not been demonstrated.

Another aspect of SPW-R’s relevant to context is the phenomenon of replay. It has been shown in rodents that the sequential firing patterns of place cells that unfold during navigation are replayed at a faster rate (Foster & Wilson, 2006; Diba & Buzsáki, 2007; Davidson, Kloosterman, & Wilson, 2009) during SPW-R’s. Critical for jumping back in time, the replayed sequence can come from a “remote” location, allowing non-recent/nearby spatial representations to be replayed (Karlsson & Frank, 2009; Xu, Baracska, O’Neill, & Csicsvari, 2019). Further, disruptions of neural activity during SPW-R’s impairs spatial memory (Girardeau, Benchenane, Wiener,

Buzsáki, & Zugaro, 2009; Ego-Stengel & Wilson, 2010; Gridchyn, Schoenenberger, O’Neill, & Csicsvari, 2020). Recent work has demonstrated that place cells phase precess during ripples (Bush, Olafsdottir, Barry, & Burgess, 2022), firing at earlier phases in the oscillation as the ripple unfolds. This precession is similar to theta phase precession reported for place cells (O’Keefe & Recce, 1993; Hafting, Fyhn, Bonnevie, Moser, & Moser, 2008) and time cells (Pastalkova et al., 2008; Ning et al., 2022), but precession during ripples unfolds faster than theta phase precession. Recent work has shown preliminary evidence for replay in humans using iEEG (Zhang, Fell, & Axmacher, 2018), MEG (Liu et al., 2019), and single-unit recordings (Vaz, Wittig Jr, Inati, & Zaghoul, 2020). Vaz et al. (2020) reported sequences during PBE’s in the anterior temporal lobe while surgical patients performed a paired associates task. Critically, the more the sequence at encoding matched the sequence during retrieval; the more likely participants would recall the correct word, but only for PBE’s coupled to MTL ripples.

The relationship between sequences during PBE’s and non-spatial sequences over slower time scales remains an outstanding question. If SPW-R’s indicate the re-statement of spatiotemporal context, and time cells represent temporal context, then one would expect to see the replay of time cell sequences during SPW-R’s or PBE’s. As far as the author knows, this has yet to be reported. Existing hippocampal time cell datasets involving a memory task (e.g., Cruzado et al., 2020) could be leveraged to test this hypothesis. Mapping the sequences reported by Vaz et al. (2020) during PBE’s onto slower sequences is less straightforward, as the “context” these sequences represent is not as clear. However, carefully examining spiking during periods without PBE’s could reveal what these sequences represent.

5.2.3 Drift at multiple time scales in the human anterior temporal lobe

Retrieved context models predict that context changes slowly over time. Indeed, neural activity drifts in the hippocampus and other brain regions (Hyman, Ma, Balaguer-Ballester, Durstewitz, & Seamans, 2012; Yaffe et al., 2014; Mankin, Diehl, Sparks, Leutgeb, & Leutgeb, 2015; Cai et al., 2016; Rashid et al., 2016; Folkerts et al., 2018; Liu et al., 2022). A scale-invariant retrieved-context model predicts that drift should occur over a wide range of time scales. Consistent with this view, Yaffe et al. (2014) reported drift in oscillatory power during encoding over seconds and over a few minutes in the human anterior temporal lobe as patients completed a paired associates task. Preliminary analyses of single units from the same task in the same brain region indicate simultaneous drift within a trial, within a list, and within a session. Consistent with findings in mouse hippocampus (Liu et al., 2022), this drift is not stochastic. The population activity of single units in the anterior lobe at equivalent points within a trial or list are similar, even when separated in time. This multiscale drift could allow one to account for contiguity and recency effects seen in some list learning experiments (Howard et al., 2008; Unsworth, 2008), as it enables the formation of associations between items within a list as well as across lists simultaneously.

References

- Adler, A., Katabi, S., Finkes, I., Israel, Z., Prut, Y., & Bergman, H. (2012). Temporal convergence of dynamic cell assemblies in the striato-pallidal network. *Journal of Neuroscience*, *32*(7), 2473-84. doi: 10.1523/JNEUROSCI.4830-11.2012
- Akhlaghpour, H., Wiskerke, J., Choi, J. Y., Taliaferro, J. P., Au, J., & Witten, I. (2016). Dissociated sequential activity and stimulus encoding in the dorsomedial striatum during spatial working memory. *eLife*, *5*, e19507.
- Anderson, J. A. (1973). A theory for the recognition of items from short memorized lists. *Psychological Review*, *80*, 417-438.
- Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, *38*, 341-380.
- Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*(2), 97-123.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, p. 89-105). New York: Academic Press.
- Averell, L., Prince, M., & Heathcote, A. (2016). Fundamental causes of systematic and random variability in recognition memory. *Journal of Memory and Language*, *88*, 51-69.
- Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*, 173-189.
- Boccaro, C. N., Sargolini, F., Thoresen, V. H., Solstad, T., Witter, M. P., Moser, E. I., & Moser, M. B. (2010). Grid cells in pre- and parasubiculum. *Nature Neuroscience*, *13*(8), 987-94.
- Bolkan, S. S., Stujenske, J. M., Parnaudeau, S., Spellman, T. J., Rauffenbart, C., Abbas, A. I., ... Kellendonk, C. (2017). Thalamic projections sustain prefrontal activity during working memory maintenance. *Nature Neuroscience*, *20*(7), 987-996. Retrieved from <http://dx.doi.org/10.1038/nn.4568>

- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329.
- Bragin, A., Engel Jr, J., Wilson, C. L., Fried, I., & Buzsáki, G. (1999). High-frequency oscillations in human brain. *Hippocampus*, *9*(2), 137–142.
- Bright, I. M., Meister, M. L., Cruzado, N. A., Tiganj, Z., Buffalo, E. A., & Howard, M. W. (2020). A temporal record of the past with a spectrum of time constants in the monkey entorhinal cortex. *Proceedings of the National Academy of Sciences*, *117*(33), 20274–20283.
- Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–76.
- Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, *112*(1), 117–28. doi: 10.1037/0033-295X.112.1.117
- Burgess, N., & O’Keefe, J. (1996). Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus*, *6*(6), 749–62.
- Bush, D., Olafsdottir, H. F., Barry, C., & Burgess, N. (2022). Ripple band phase precession of place cell firing during replay. *Current Biology*, *32*(1), 64–73.
- Buzsáki, G. (1986). Hippocampal sharp waves: their origin and significance. *Brain research*, *398*(2), 242–252.
- Buzsáki, G. (1996). The hippocampo-neocortical dialogue. *Cerebral Cortex*, *6*, 81–92.
- Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, *25*(10), 1073–1188.
- Buzsáki, G., Leung, L. S., & Vanderwolf, C. H. (1983). Cellular bases of hippocampal eeg in the behaving rat. *Brain Research Review*, *6*, 139–171.
- Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron*, *80*(3), 751–764.
- Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., ... Silva, A. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, *534*(7605), 115–118.
- Cao, R., Bladon, J. H., Charczynski, S. J., Hasselmo, M. E., & Howard, M. W. (2021). Internally generated time in the rodent hippocampus is logarithmically compressed. *bioRxiv*.

- Chater, N., & Brown, G. D. A. (2008). From universal laws of cognition to specific cognitive models. *Cognitive Science*, *32*(1), 36-67. doi: 10.1080/03640210701801941
- Chen, Y. Y., Aponik-Gremillion, L., Bartoli, E., Yoshor, D., Sheth, S. A., & Foster, B. L. (2021). Stability of ripple events during task engagement in human hippocampus. *Cell reports*, *35*(13), 109304.
- Cruzado, N. A., Tiganj, Z., Brincat, S. L., Miller, E. K., & Howard, M. W. (2020). Conjunctive representation of what and when in monkey hippocampus and lateral prefrontal cortex during an associative memory task. *Hippocampus*, *30*(12), 1332–1346.
- da Costa Pinto, A., & Baddeley, A. (1991). Where did you park your car? analysis of a naturalistic long-term recency effect. *European Journal of Cognitive Psychology*, *3*(3), 297–313.
- Dannenberg, H., Kelley, C., Hoyland, A., Monaghan, C. K., & Hasselmo, M. E. (2019). The firing rate speed code of entorhinal speed cells differs across behaviorally relevant time scales and does not depend on medial septum inputs. *Journal of Neuroscience*, 1450–18.
- Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review*, *112*(1), 3-42.
- Davidson, T. J., Kloosterman, F., & Wilson, M. A. (2009). Hippocampal replay of extended experience. *Neuron*, *63*(4), 497-507. doi: 10.1016/j.neuron.2009.07.027
- Davis, O., Rizzuto, D., Geller, A. S., & Kahana, M. J. (2008). Temporal associative processes revealed by intrusions in paired-associate recall. *Psychonomic Bulletin & Review*, *15*(1), 64-9.
- Dede, A. J., Frascino, J. C., Wixted, J. T., & Squire, L. R. (2016). Learning and remembering real-world events after medial temporal lobe damage. *Proceedings of the National Academy of Sciences*, *113*(47), 13480–13485.
- Deshmukh, S. S., & Knierim, J. J. (2011). Representation of non-spatial and spatial information in the lateral entorhinal cortex. *Frontiers in behavioral neuroscience*, *5*, 69.
- Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, *10*(10), 1241-2.
- Donchin, E. (1981). Surprise!... surprise? *Psychophysiology*, *18*(5), 493–513.

- Donkin, C., & Nosofsky, R. M. (2012). A power-law model of psychological memory strength in short- and long-term recognition. *Psychological Science*. doi: 10.1177/0956797611430961
- Egorov, A. V., Hamam, B. N., Fransén, E., Hasselmo, M. E., & Alonso, A. A. (2002). Graded persistent activity in entorhinal cortex neurons. *Nature*, *420*(6912), 173-8.
- Ego-Stengel, V., & Wilson, M. A. (2010). Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. *Hippocampus*, *20*(1), 1-10.
- Eichenbaum, H. (2017). On the integration of space, time, and memory. *Neuron*, *95*(5), 1007-1018. doi: 10.1016/j.neuron.2017.06.036
- Eichenbaum, H., Yonelinas, A., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123-152.
- Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145-154.
- Fechner, G. (1860/1912). *Elements of psychophysics*. Vol. I. Houghton Mifflin.
- Folkerts, S., Rutishauser, U., & Howard, M. (2018). Human episodic memory retrieval is accompanied by a neural contiguity effect. *Journal of Neuroscience*, *38*, 4200-4211.
- Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, *440*(7084), 680-3.
- Frank, L. M., Brown, E. N., & Wilson, M. (2000). Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron*, *27*(1), 169-178.
- Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods and Instrumentation*, *14*, 375-399.
- Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., & Moser, M. B. (2004). Spatial representation in the entorhinal cortex. *Science*, *305*(5688), 1258-64.
- Geller, A. S., Schlefer, I. K., Sederberg, P. B., Jacobs, J., & Kahana, M. J. (2007). PyEPL: a cross-platform experiment-programming library. *Behavior Research Methods*, *39*(4), 950-8.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1-67.

- Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., & Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature neuroscience*, *12*(10), 1222–1223.
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, *5*, 351–360.
- Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., & Gretz, A. L. (1980). A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 355–369.
- Gridchyn, I., Schoenenberger, P., O’Neill, J., & Csicsvari, J. (2020). Assembly-specific disruption of hippocampal replay leads to selective memory deficit. *Neuron*, *106*(2), 291–300.
- Guderian, S., Schott, B. H., Richardson-Klavehn, A., & Düzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proceedings of the National Academy of Sciences*, *106*(13), 5365–5370.
- Guo, C., Huson, V., Macosko, E. Z., & Regehr, W. G. (2021). Graded heterogeneity of metabotropic signaling underlies a continuum of cell-intrinsic temporal responses in unipolar brush cells. *Nature Communications*, *12*(1), 5491.
- Hacker, M. J. (1980). Speed and accuracy of recency judgments for events in short-term memory. *Journal of Experimental Psychology: Human Learning and Memory*, *15*, 846–858.
- Hafting, T., Fyhn, M., Bonnevie, T., Moser, M. B., & Moser, E. I. (2008). Hippocampus-independent phase precession in entorhinal grid cells. *Nature*, *453*(7199), 1248–52.
- Hahn, T. T., McFarland, J. M., Berberich, S., Sakmann, B., & Mehta, M. R. (2012). Spontaneous persistent activity in entorhinal cortex modulates cortico-hippocampal interaction in vivo. *Nature neuroscience*, *15*(11), 1531.
- Hampson, R. E., & Deadwyler, S. A. (2003). Temporal firing characteristics and the strategic role of subicular neurons in short-term memory. *Hippocampus*, *13*(4), 529–541.
- Hasselmo, M. E. (2012). *How we remember: Brain mechanisms of episodic memory*. Cambridge, MA: MIT Press.
- Hasselmo, M. E., Bodelón, C., & Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: Separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Computation*, *14*, 793–817.

- Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General*, *143*(2), 575.
- Healey, M. K., Long, N. M., & Kahana, M. J. (2019). Contiguity in episodic memory. *Psychonomic bulletin & review*, *26*(3), 699–720.
- Henke, J., Bunk, D., von Werder, D., Häusler, S., Flanagin, V. L., & Thurley, K. (2021). Distributed coding of duration in rodent prefrontal cortex during time reproduction. *Elife*, *10*, e71612.
- Hintzman, D. (1987). Recognition and recall in minerva 2: Analysis of the ‘recognition-failure’ paradigm. In P. Morris (Ed.), *Modelling cognition* (p. 215-229). New York: Wiley.
- Hintzman, D. L. (1969). Recognition time: Effects of recency, frequency and the spacing of repetitions. *Journal of Experimental Psychology*, *79*(1p1), 192.
- Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition*, *38*(1), 102-15.
- Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgements: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, *33*, 1-18.
- Hockley, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 497-512.
- Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 598-615.
- Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition*, *3*(2), 197-209.
- Howard, M. W. (2018). Memory as perception of the past: Compressed time in mind and brain. *Trends in Cognitive Sciences*, *22*, 124-136.
- Howard, M. W. (2022). Formal models of memory based on temporally-varying representations. *arXiv preprint arXiv:2201.01796*.
- Howard, M. W., & Eichenbaum, H. (2015). Time and space in the hippocampus. *Brain Research*, *1621*, 345-354.
- Howard, M. W., Esfahani, Z. G., Le, B., & Sederberg, P. B. (2023). Foundations of a temporal rl. *arXiv preprint arXiv:2302.10163*.

- Howard, M. W., & Hasselmo, M. E. (2020). Cognitive computation using neural representations of time and space in the laplace domain. *arXiv preprint arXiv:2003.11668*.
- Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 923-941.
- Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*(1), 85-98.
- Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of Neuroscience*, *34*(13), 4692-707. doi: 10.1523/JNEUROSCI.5808-12.2014
- Howard, M. W., Shankar, K. H., Aue, W., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, *122*(1), 24-53.
- Howard, M. W., Viskontas, I. V., Shankar, K. H., & Fried, I. (2012). Ensembles of human MTL neurons “jump back in time” in response to a repeated stimulus. *Hippocampus*, *22*(9), 1833-1847.
- Howard, M. W., Youker, T. E., & Venkatadass, V. (2008). The persistence of memory: Contiguity effects across several minutes. *Psychonomic Bulletin & Review*, *15*(PMC2493616), 58-63.
- Høydal, Ø. A., Skytøen, E. R., Andersson, S. O., Moser, M.-B., & Moser, E. I. (2019). Object-vector coding in the medial entorhinal cortex. *Nature*, *568*(7752), 400.
- Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, *81*(5), 1165–1178.
- Husserl, E. (1966). *The phenomenology of internal time-consciousness*. Bloomington, IN: Indiana University Press.
- Hyde, R. A., & Strowbridge, B. W. (2012). Mnemonic representations of transient stimuli and temporal sequences in the rodent hippocampus in vitro. *Nature Neuroscience*, *15*(10), 1430-8. doi: 10.1038/nn.3208
- Hyman, J. M., Ma, L., Balaguer-Ballester, E., Durstewitz, D., & Seamans, J. K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proceedings of the National Academy of Sciences USA*, *109*, 5086-91. doi: 10.1073/pnas.1114415109

- Inagaki, H. K., Inagaki, M., Romani, S., & Svoboda, K. (2018). Low-dimensional and monotonic preparatory activity in mouse anterior lateral motor cortex. *Journal of Neuroscience*, *38*(17), 4163–4185.
- James, W. (1890). *The principles of psychology*. New York: Holt.
- Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences*, *106*(45), 19156–19161.
- Jutras, M. J., & Buffalo, E. A. (2010). Recognition memory signals in the macaque hippocampus. *Proceedings of the National Academy of Sciences*, *107*(1), 401–406.
- Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*, 103-109.
- Kahana, M. J. (2012). *Foundations of human memory*. OUP USA.
- Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, *30*(6), 841-9.
- Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag-recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 530-540.
- Karlsson, M. P., & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience*, *12*(7), 913-8. doi: 10.1038/nn.2344
- Kellen, D., & Klauer, K. C. (2014). Discrete-state and continuous models of recognition memory: Testing core properties under minimal assumptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1795–1804.
- Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. *Nature*, *491*(7426), 761-4. doi: 10.1038/nature11587
- Killian, N. J., Potter, S. M., & Buffalo, E. A. (2015). Saccade direction encoding in the primate entorhinal cortex during visual exploration. *Proceedings of the National Academy of Sciences USA*, *112*(51), 15743-8. doi: 10.1073/pnas.1417059112
- Kim, J., Ghim, J.-W., Lee, J. H., & Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *Journal of Neuroscience*, *33*(34), 13834-47. doi: 10.1523/JNEUROSCI.1443-13.2013
- Klink, R., & Alonso, A. (1997). Muscarinic modulation of the oscillatory and repetitive firing properties of entorhinal cortex layer ii neurons. *Journal of neurophysiology*, *77*(4), 1813–1828.

- Kraus, B. J., Brandon, M. P., Robinson, R. J., Connerney, M. A., Hasselmo, M. E., & Eichenbaum, H. (2015). During running in place, grid cells integrate elapsed time and distance run. *Neuron*, *88*(3), 578–589.
- Kraus, B. J., Robinson, R. J., 2nd, White, J. A., Eichenbaum, H., & Hasselmo, M. E. (2013). Hippocampal "time cells": time versus path integration. *Neuron*, *78*(6), 1090–101. doi: 10.1016/j.neuron.2013.04.015
- Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology*, *34*(5/6), 419–426.
- Lebedev, M. A., O'Doherty, J. E., & Nicolelis, M. A. (2008). Decoding of temporal intervals from cortical ensemble activity. *Journal of neurophysiology*, *99*(1), 166–186.
- Leitner, F. C., Melzer, S., Lütcke, H., Pinna, R., Seeburg, P. H., Helmchen, F., & Monyer, H. (2016). Spatially segregated feedforward and feedback neurons support differential odor processing in the lateral entorhinal cortex. *Nature Neuroscience*.
- Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, *38*(2), 317–327.
- Le Van Quyen, M., Staba, R., Bragin, A., Dickson, C., Valderrama, M., Fried, I., & Engel, J. (2010). Large-scale microelectrode recordings of high-frequency gamma oscillations in human cortex during sleep. *Journal of Neuroscience*, *30*(23), 7770–7782.
- Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., & Burgess, N. (2009). Boundary vector cells in the subiculum of the hippocampal formation. *Journal of Neuroscience*, *29*(31), 9771–7.
- Liu, Y., Levy, S., Mau, W., Geva, N., Rubin, A., Ziv, Y., ... Howard, M. (2022). Consistent population activity on the scale of minutes in the mouse hippocampus. *Hippocampus*, *32*(5), 359–372.
- Liu, Y., Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2019). A neural microcircuit model for a scalable scale-invariant representation of time. *Hippocampus*.
- Logan, G. D. (2018). Automatic control: How experts act without thinking. *Psychological Review*, *125*(4), 453.
- Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337–63. doi: 10.1037/a0039036

- MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories immobilized rats. *Journal of Neuroscience*, *33*(36), 14607–14616.
- MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontinuous events. *Neuron*, *71*(4), 737–749.
- Machens, C. K., Romo, R., & Brody, C. D. (2010). Functional, but not anatomical, separation of what and when in prefrontal cortex. *Journal of Neuroscience*, *30*(1), 350–360.
- Malmberg, K. J., & Annis, J. (2012). On the relationship between memory and perception: Sequential dependencies in recognition memory testing. *J Exp Psychol Gen*, *141*(2), 233–59. doi: 10.1037/a0025277
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271.
- Mankin, E. A., Diehl, G. W., Sparks, F. T., Leutgeb, S., & Leutgeb, J. K. (2015). Hippocampal CA2 activity patterns change over time to a larger extent than between spatial contexts. *Neuron*, *85*(1), 190–201. doi: 10.1016/j.neuron.2014.12.001
- Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, *109*, 19462–7. doi: 10.1073/pnas.1214107109
- Manning, J. R., Polyn, S. M., Litt, B., Baltuch, G., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Science, USA*, *108*(31), 12893–7.
- Manns, J. R., Howard, M. W., & Eichenbaum, H. B. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*(3), 530–540.
- Mau, W., Sullivan, D. W., Kinsky, N. R., Hasselmo, M. E., Howard, M. W., & Eichenbaum, H. (2018). The same hippocampal CA1 population simultaneously codes temporal information over multiple timescales. *Current Biology*, *28*, 1499–1508.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346–373.

- McElree, B., & Doshier, B. A. (1993). Serial recovery processes in the recovery of order information. *Journal of Experimental Psychology: General*, *122*, 291-315.
- Meister, M. L., & Buffalo, E. A. (2018). Neurons in primate entorhinal cortex represent gaze position in multiple spatial reference frames. *Journal of Neuroscience*, *38*(10), 2430-2441.
- Mello, G. B., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology*, *25*(9), 1113-1122.
- Mensink, G.-J. M., & Raaijmakers, J. G. W. (1989). A model for contextual fluctuation. *Journal of Mathematical Psychology*, *33*, 172-186.
- Meyer, T., & Rust, N. C. (2018). Single-exposure visual memory judgments are reflected in inferotemporal cortex. *Elife*, *7*. doi: 10.7554/eLife.32259
- Milner, B. (1959). The memory defect in bilateral hippocampal lesions. *Psychiatric research reports*, *11*, 43.
- Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *10*, 465-501.
- Moreton, B. J., & Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin & Review*, *17*, 510-515.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61-64.
- Mormann, F., Kornblith, S., Quiroga, R. Q., Kraskov, A., Cerf, M., Fried, I., & Koch, C. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *Journal of Neuroscience*, *28*(36), 8865-8872.
- Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482-488.
- Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609-626.
- Murdock, B. B., & Anderson, R. E. (1975). Encoding, storage and retrieval of item information. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (p. 145-194). Hillsdale, New Jersey: Erlbaum.
- Murdock, B. B., & Dufty, P. O. (1972). Strength theory and recognition memory. *Journal of Experimental Psychology*, *94*, 284-290.

- Murdock, B. B., Jr. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology*, *60*, 222-34.
- Naya, Y., Chen, H., Yang, C., & Suzuki, W. A. (2017). Contributions of primate prefrontal cortex and medial temporal lobe to temporal-order memory. *Proceedings of the National Academy of Sciences*, *114*(51), 13555–13560.
- Naya, Y., & Suzuki, W. (2011). Integrating what and when across the primate medial temporal lobe. *Science*, *333*(6043), 773-776.
- Ning, W., Bladon, J. H., & Hasselmo, M. E. (2022). Complementary representations of time in the prefrontal cortex and hippocampus. *Hippocampus*, *32*(8), 577–596.
- Norman, D. A., & Wickelgren, W. A. (1969). Strength theory of decision rules and latency in short-term memory. *Journal of Mathematical Psychology*, *6*, 192-208.
- Norman, Y., Yeagle, E. M., Khuvis, S., Harel, M., Mehta, A. D., & Malach, R. (2019). Hippocampal sharp-wave ripples linked to visual episodic recollection in humans. *Science*, *365*(6454), eaax1030.
- Nosofsky, R. M., Cox, G. E., Cao, R., & Shiffrin, R. M. (2014). An exemplar-familiarity model predicts short-term and long-term probe recognition across diverse forms of memory search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1524.
- Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, *118*(2), 280-315.
- Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1023-35. doi: 10.1016/j.neubiorev.2009.12.014
- Okada, R. (1971). Decision latencies in short-term recognition memory. *Journal of Experimental Psychology*, *90*(1), 27.
- O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental neurology*, *51*(1), 78–109.
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171-175.
- O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*(3), 317-30.

- Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology: General*, *139*(2), 341-64.
- Palombo, D. J., Di Lascio, J. M., Howard, M. W., & Verfaellie, M. (2019). Medial temporal lobe amnesia is associated with a deficit in recovering temporal context. *Journal of cognitive neuroscience*, *31*(2), 236–248.
- Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*(5894), 1322-7.
- Pilkiw, M., Insel, N., Cui, Y., Finney, C., Morrissey, M. D., & Takehara-Nishiuchi, K. (2017). Phasic and tonic neuron ensemble codes for stimulus-environment conjunctions in the lateral entorhinal cortex. *Elife*, *6*, e28611.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2021). nlme: Linear and nonlinear mixed effects models [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=nlme> (R package version 3.1-153)
- Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*, 129-156.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of experimental psychology: human learning and memory*, *2*(5), 509.
- Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of experimental psychology*, *81*(1), 10.
- Province, J. M., & Rouder, J. N. (2012). Evidence for discrete-state processing in recognition memory. *Proceedings of the National Academy of Sciences USA*, *109*(36), 14357-62. doi: 10.1073/pnas.1103880109
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, p. 207-262). New York: Academic Press.
- Rashid, A. J., Yan, C., Mercaldo, V., Hsiang, H.-L. L., Park, S., Cole, C. J., ... others (2016). Competition between engrams influences fear memory formation and recall. *Science*, *353*(6297), 383–387.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.
- Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*(3), 190-214.

- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological review*, *111*(2), 333.
- Reddy, L., Zoefel, B., Possel, J. K., Peters, J., Dijksterhuis, D. E., Poncet, M., . . . Self, M. W. (2021). Human hippocampal neurons track moments in a sequence of events. *Journal of Neuroscience*, *41*(31), 6714–6725.
- Rossi-Pool, R., Salinas, E., Zainos, A., Alvarez, M., Vergara, J., Parga, N., & Romo, R. (2016). Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. *Proceedings of the National Academy of Sciences*, *113*(49), E7966–E7975.
- Rossi-Pool, R., Zizumbo, J., Alvarez, M., Vergara, J., Zainos, A., & Romo, R. (2019). Temporal signals underlying a cognitive process in the dorsal premotor cortex. *Proceedings of the National Academy of Sciences*, *116*(15), 7523–7532.
- Rotello, C. M. (2017). Cognitive psychology of memory. In J. T. Wixted (Ed.), (2nd edition ed., Vol. 2, p. 201-225). Elsevier.
- Rubin, D. C., Hinton, S., & Wenzel, A. (1999). The precise time course of retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1161.
- Sakon, J. J., & Kahana, M. J. (2022). Hippocampal ripples signal contextually mediated episodic recall. *Proceedings of the National Academy of Sciences*, *119*(40), e2201657119.
- Salz, D. M., Tiganj, Z., Khasnabish, S., Kohley, A., Sheehan, D., Howard, M. W., & Eichenbaum, H. (2016). Time cells in hippocampal area CA3. *Journal of Neuroscience*, *36*, 7476-7484.
- Schonhaut, D. R., Aghajani, Z. M., Kahana, M. J., & Fried, I. (2022). A neural code for spatiotemporal context. *bioRxiv*, 2022–05.
- Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Research*, *20*(8), 645-69.
- Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, *16*(11), 898-904.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*, 11-21.
- Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*, 893-912.

- Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., & Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience*, *23*(34), 10809-14.
- Shankar, K. H., & Howard, M. W. (2012). A scale-invariant internal representation of time. *Neural Computation*, *24*(1), 134-193.
- Shankar, K. H., & Howard, M. W. (2013). Optimally fuzzy temporal memory. *Journal of Machine Learning Research*, *14*, 3753-3780.
- Shepard, R. N., & Teghtsoonian, M. (1961). Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology*, *62*, 302-309.
- Shiffrin, R. M., Ratcliff, R., Murnane, K., & Nobel, P. (1993). Todam and the list-strength and list-length effects: A reply to Murdock and Kahana. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1445-1449.
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM — retrieving effectively from memory. *Psychonomic Bulletin and Review*, *4*, 145-166.
- Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological review*, *116*(2), 283.
- Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition*, *40*(2), 177-90. doi: 10.3758/s13421-011-0142-8
- Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, *27*, 279-306.
- Standing, L. (1973). Learning 10000 pictures. *The Quarterly journal of experimental psychology*, *25*(2), 207-222.
- Sternberg, S. (2016). In defence of high-speed memory scanning. *Quarterly Journal of Experimental Psychology*, *69*(10), 2020-2075.
- Suzuki, W. A., Miller, E. K., & Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. *Journal of Neurophysiology*, *78*(2), 1062-1081.
- Tahvildari, B., Fransén, E., Alonso, A. A., & Hasselmo, M. E. (2007). Switching between "On" and "Off" states of persistent activity in lateral entorhinal layer III neurons. *Hippocampus*, *17*(4), 257-63.

- Takeuchi, D., Hirabayashi, T., Tamura, K., & Miyashita, Y. (2011, Mar). Reversal of interlaminar signal between sensory and memory processing in monkey temporal cortex. *Science*, *331*(6023), 1443-7. doi: 10.1126/science.1199967
- Taxidis, J., Pnevmatikakis, E., Mylavarapu, A. L., Arora, J. S., Samadian, K. D., Hoffberg, E. A., & Golshani, P. (2018). Emergence of stable sensory and dynamic temporal representations in the hippocampus during working memory. *BioRxiv*, 474510.
- Terada, S., Sakurai, Y., Nakahara, H., & Fujisawa, S. (2017). Temporal and rate coding for discrete event sequences in the hippocampus. *Neuron*, *94*, 1-15.
- Tiganj, Z., Cromer, J. A., Roy, J. E., Miller, E. K., & Howard, M. W. (2017). Compressed timeline of recent experience in monkey LPFC. *bioRxiv*, 126219.
- Tiganj, Z., Cromer, J. A., Roy, J. E., Miller, E. K., & Howard, M. W. (2018). Compressed timeline of recent experience in monkey LPFC. *Journal of Cognitive Neuroscience*, *30*, 935-950.
- Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2015). A simple biophysically plausible model for long time constants in single neurons. *Hippocampus*, *25*(1), 27-37.
- Tiganj, Z., Kim, J., Jung, M. W., & Howard, M. W. (2017). Sequential firing codes for time in rodent mPFC. *Cerebral Cortex*, *27*, 5663-5671.
- Tiganj, Z., Shankar, K. H., & Howard, M. W. (2017). Scale invariant value computation for reinforcement learning in continuous time. In *AAAI 2017 spring symposium series - science of intelligence: Computational principles of natural and artificial intelligence*.
- Tiganj, Z., Singh, I., Esfahani, Z. G., & Howard, M. W. (2021). Scanning a compressed ordered representation of the future. *bioRxiv*, 229617.
- Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B., & Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature*, *561*, 57-62. Retrieved from <https://doi.org/10.1038/s41586-018-0459-6> doi: 10.1038/s41586-018-0459-6
- Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*(1), 1-12.
- Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, *21*, 437-484.

- Uitvlugt, M. G., & Healey, M. K. (2019). Temporal proximity links unrelated news events in memory. *Psychological science*, *30*(1), 92–104.
- Unsworth, N. (2008). Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. *Journal of Memory and Language*, *59*, 223–236.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–92.
- Vanderwolf, C. H. (1969). Hippocampal electrical activity and voluntary movement of the rat. *Electroencephalography and Clinical Neurophysiology*, *26*, 407–418.
- Van Essen, D. C., Newsome, W. T., & Maunsell, J. H. (1984). The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Research*, *24*(5), 429–48.
- Vargas, R., Thorsteinsson, H., & Karlsson, K. (2012). Spontaneous neural activity of the anterodorsal lobe and entopeduncular nucleus in adult zebrafish: a putative homologue of hippocampal sharp waves. *Behavioural brain research*, *229*(1), 10–20.
- Vaz, A. P., Inati, S. K., Brunel, N., & Zaghoul, K. A. (2019). Coupled ripple oscillations between the medial temporal lobe and neocortex retrieve human memory. *Science*, *363*(6430), 975–978.
- Vaz, A. P., Wittig Jr, J. H., Inati, S. K., & Zaghoul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. *Science*, *367*(6482), 1131–1134.
- Wang, C., Chen, X., Lee, H., Deshmukh, S. S., Yoganarasimha, D., Savelli, F., & Knierim, J. J. (2018). Egocentric coding of external items in the lateral entorhinal cortex. *Science*, *362*(6417), 945–949.
- Wilson, M. A., & McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, *261*, 1055–8.
- Witter, M. P., Wouterlood, F. G., Naber, P. A., & Van Haeften, T. (2000). Anatomical organization of the parahippocampal-hippocampal network. *Annals of the New York Academy of Science*, 1–24.
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–76.
- Wood, E. R., Dudchenko, P. A., Robitsek, R. J., & Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron*, *27*(3), 623–33.

- Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 787-807.
- Xiang, J.-Z., & Brown, M. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology*, *37*(4-5), 657-676.
- Xu, H., Baracska, P., O'Neill, J., & Csicsvari, J. (2019). Assembly responses of hippocampal ca1 place cells predict learned behavior in goal-directed spatial tasks on the radial eight-arm maze. *Neuron*, *101*(1), 119-132.
- Yaffe, R. B., Kerr, M. S. D., Damera, S., Sarma, S. V., Inati, S. K., & Zaghoul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences*, *111*(52), 18727-32. doi: 10.1073/pnas.1417017112
- Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*, 747-763.
- Yoshida, M., Fransén, E., & Hasselmo, M. E. (2008). mGluR-dependent persistent firing in entorhinal cortex layer III neurons. *European Journal of Neuroscience*, *28*(6), 1116-26.
- Zhang, H., Fell, J., & Axmacher, N. (2018). Electrophysiological mechanisms of human memory consolidation. *Nature Communications*, *9*(1), 4103.
- Zhou, J., Osth, A. F., & Smith, P. L. (2023). The spatiotemporal gradient of intrusion errors in continuous outcome source memory: Source retrieval is affected by both guessing and intrusions. *Cognitive Psychology*, *141*, 101552.

CURRICULUM VITAE

