Comparing Impact of Time Lag and Item Lag in Relative Judgment of Recency

Sahaj Singh Maini (sahmaini@iu.edu), Louis Labuzienski (llabuzie@iu.edu),
Saurabh Gulati (sagulati@iu.edu), Zoran Tiganj (ztiganj@iu.edu)
Department of Computer Science
Luddy School of Informatics, Computing, and Engineering
Indiana University Bloomington

Abstract

Many memory models suggest self-terminating backward scanning along a memory representation. In these models, time to retrieve a particular item from memory could depend on how far in the past the item was presented or on the number of items presented since that item. To investigate which of these two types of memory representation is more likely, we designed a relative Judgment of Recency (JOR) task with variable presentation rates. The variable presentation rate deconfounded the age of memory and the number of intervening items. Our results favor the hypothesis that memory representation is temporally organized. This result is important for advancing memory models and for building stronger ties between cognitive and neural models of memory.

Keywords: Judgment of Recency; Variable presentation time; Time versus items.

Introduction

In the relative judgment of recency (JOR) task, participants are presented with a sequence of items followed by a probe consisting of two items from the sequence. The participants have to select the item that was presented more recently. Results from previous studies suggest that the response time (RT) grows sublinearly with the lag to the more recent probe and that it does not depend on the lag to the more distant probe (Muter, 1979; Hacker, 1980; Hockley, 1984; Tiganj, Singh, Esfahani, & Howard, in press).

The finding that RT depends only on the lag to the more recent probe was often explained through self-terminating backward scan along a memory representation (Muter, 1979; Hacker, 1980; Hockley, 1984; McElree & Dosher, 1993; Howard, 2014). The finding that RT grows sublinearly with the lag to the more recent probe was used to argue that the temporally organized memory is compressed such that the more recent past is represented with higher resolution than the more distant past (Howard, 2014). This is consistent with the knowledge that memory gets worse for events further in the past. While these models rely on backward scan, that scan could be along a temporally organized or ordinally organized memory representation.

In most JOR experiments, time lag (age of memory) and item lag (the number of intervening items) were confounded. This made it impossible to distinguish whether scanning along a temporally organized or item organized memory representation would better explain the data. Hintzman (2004) studied absolute JOR with variable presentation rate to deconfound time lag and item lag. In Hintzman (2004) subjects did absolute JOR on a long list that was made up of alternating fast and slow blocks with 25 to 50 items each. Hintzman (2004) concluded that response times were a function of time lag with no added contribution from the item lag. While this study provided important insight into recency judgments over relatively long temporal scales (tens of seconds), it remains unclear whether the same account would hold for shorter temporal scales (several seconds) that are typically used in relative JOR tasks and that are commonly modeled with different memory models, *e.g.*, Atkinson and Shiffrin (1968).

Studies of serial recall have also made an important contribution for understating the role of time in short-term memory of serial order. Separating the output interference, rather than output time, was argued to be critical in serial recall (Lewandowsky, Duncan, & Brown, 2004). Also, temporal representation was argued to be unnecessary for short-term memory for serial order (Lewandowsky, Brown, Wright, & Nimmo, 2006; Lewandowsky, Oberauer, & Brown, 2009).

Here we designed a version of relative JOR task that deconfounds item lag and time lag by varying the presentation rate within each list. A varying presentation rate was obtained by inserting "blank" letters which introduced a gap between otherwise consecutive letters, creating longer inter-stimulus intervals. This manipulation allowed us to study response time changes as a function of both item lag and time lag. Our results are consistent with findings from Hintzman (2004) and extend those findings to much shorter temporal scales proving a helpful insight for future memory models and for efforts to better understand the neural underpinnings of memory.

Materials and Methods

Participants were presented with a list of 9, 11, or 13 letters (all consonants) at the rate of 5.5 letters per second or one letter every 0.181 s. At each trial, the presentation order was random. A single letter could be presented at most once in a single list – no repetitions were allowed (Figure 1a).

To introduce variability to the presentation rate and deconfound time lag and item lag, after each letter was presented there was a 40% probability of a gap. A gap was presented for the same duration as a single letter. Only a single consecutive gap was possible. Thus duration of each list was at most 26×0.181 s $= 4.73$ s (Figure 1b).

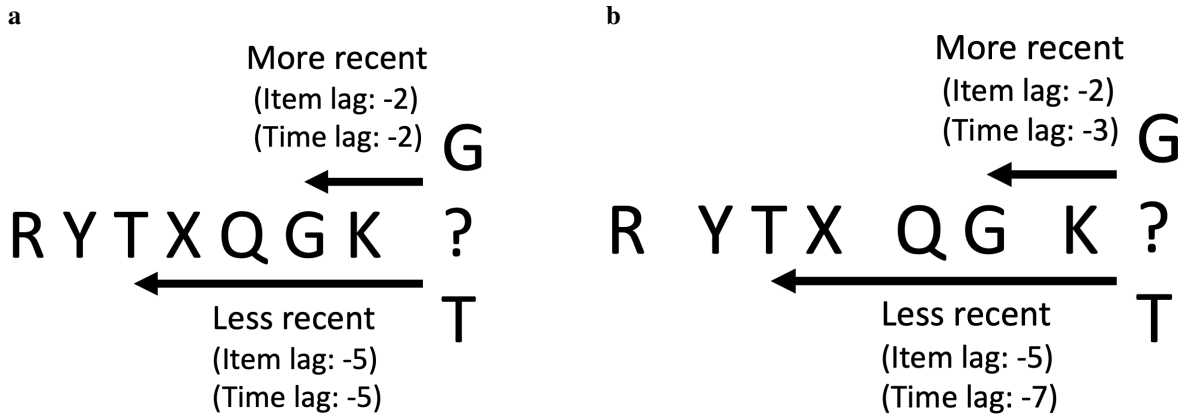At the end of the list, two of the last seven letters were chosen randomly, and the participants were asked to indicate

Figure 1: **Schematic of relative JOR task with fixed and variable presentation rate a.** JOR task with fixed presentation rate. Participants are shown a list of letters (such as RYT...) followed by a probe containing two letters from the list (here, G and T). Participants are asked to select the probe item that was experienced more recently. In this example, the probe G is the correct answer. Lag for each probe is defined as the number of steps backward in the list necessary to find the probe. In this case, item lag (the number of intervening items) and time lag (age of memory) are the same because the presentation rate was fixed. **b.** JOR task with variable presentation rate. Gaps are introduced to the list such that the duration of a single gap is equal to the duration of a single letter. Gaps occur with a probability of 40% such that at most one consecutive gap is possible. Time lag is defined as the number of steps backward in the list necessary to find the probe counting both gaps and letters, while item lag is defined as the number of steps backward in the list necessary to find the probe but counting only the letters and not the gaps.

using the left or right arrow key which of the two letters had appeared more recently. In Figure 1, G and T are presented as the probe items. Because G was presented more recently than T, the correct answer is G. If the participant did not make a response within 2 s, the trial was terminated.

The distance to the more recent probe stimulus varied from item lag -1 to -6. The item lag of the less recent probe varied from -2 to -7. This gave us 21 possible combinations of lags, which were presented in random order. Each participant completed 200 trials.

The participants were selected using the Amazon Mechanical Turk platform and they completed the task *via* an online interface designed using jsPsych (De Leeuw, 2015). Participants were compensated for their time. Each participant spent about 20 minutes performing the task. Prior to the beginning of the experiment, participants were given written instructions. They were also given three demo trials where letters were presented at a slower pace in order to illustrate the structure of the task. Participants were required to complete those three trials successfully before proceeding with the experiment. If they made an error in any of the three trials, that trial was repeated. The three demo trials were not used in the analysis.

The study materials and protocol were approved by the Institutional Review Board. A total of 33 participants signed up for the study. Two participants withdrew from the study. Data from additional two participants were excluded because their overall accuracy was no better than chance. The results below present the analysis of the data from the remaining 29 participants.

The procedure of this experiment was similar to the proce-

dure of the Experiment 2 of Hacker (1980) and JOR experiment in Tiganj et al. (in press), with the main difference in that here we used a variable presentation rate. Other than the variable presentation rate, unlike the Hacker (1980) study, in this experiment participants were never given foils that did not appear in the list and unlike in Tiganj et al. (in press) study, participants were not given the option to respond indicating that they did not remember either of the probes. Also, unlike the previous studies, this study was conducted online.

Results

Accuracy was similar to that in the previous studies

The probability that participants selected the more recent probe was $.69 \pm .02$. The accuracy was $.82 \pm .01$ when the time lag of the more recent probe was -1 and dropped to $.52 \pm .04$ when the time lag was -6. For time lags -6 and -5 the probability of choosing the more recent probe was not different from chance (for lag -6: Chi-squared prop test, $\chi^2(29) = 37.24$, p-value=.14; for time lag -5: Chi-squared prop test, $\chi^2(29) = 42.55$, p-value=.05 and accuracy was $.57 \pm .02$). Time lag -4 had an accuracy of 0.59 ± 0.02 and was significantly higher than chance (Chi-squared prop test, $\chi^2(29) = 64.10$, $p < 0.01$). These values are similar to those in previous studies (Tiganj et al., in press). Figure 2 shows accuracy as a function of time lag and item lag to the less and to the more recent probe.

RT was better explained by time lag than by item lag

The main objective of our analysis was to evaluate whether RTs are better explained by time lag (the age of the memory)
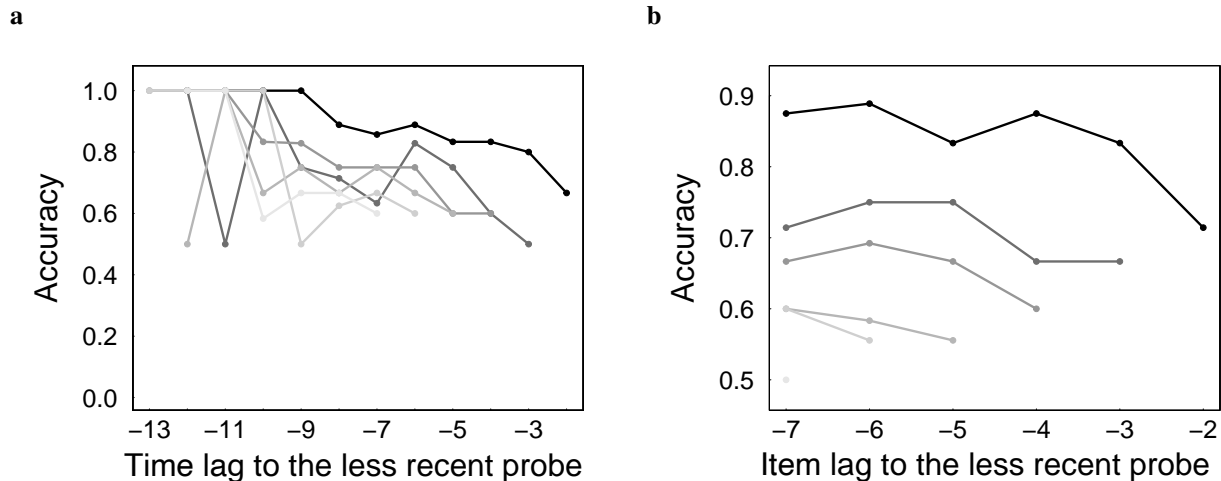
Figure 2: **Accuracy in JOR as a function of more and less recent probe. a.** Shades of gray represent time lag to the more recent probe such that the darkest shade corresponds to lag of -1 and the lightest to the lag of -11. Time on the x-axis is in units of seconds with a single time lag corresponding to item presentation time of 0.181 s. Therefore the entire x-axis corresponds to 1.81 s. **b.** Shades of gray represent item lag to the more recent probe such that the darkest shade corresponds to lag of -1 and the lightest to the lag of -6.

or item lag (the number of intervening items). Given that previous studies have shown that RT in absolute JOR does not depend on the lag to the more distant probe, we focused our analysis on the lag to the more recent probe.

Figure 3a shows RT as a function of time to the more recent probe, grouped by item lags to the more recent probe (shown in different shades). Figure 3b shows RT as a function of item lag to the more recent probe for different times to the more recent probe. Visual inspection of these two figures suggests that lines in Figure 3a are not parallel to the x-axis, while in Figure 3b they are parallel to the x-axis. If that is the case, it will support the hypothesis that time lag explains RT better than item lag.

Statistical analysis confirmed these visual impressions. Specifically, we conducted two different analyses, one based on linear mixed-effects models and the other one based on a Bayesian t-test of slopes. We compared two linear mixed-effects models in predicting RT. In the first model, item lag to the more recent probe was treated as a random effect (i.e., allowing independent intercept for each lag) and within-lag, time to the more recent probe was found to be a significant fixed effect ($.044 \pm .007$ s, $t(14) = 6.54$, $p < 0.001$). In contrast, in the second model, time to the more recent probe was a random effect and within-time, the fixed effect of item lag to the more recent probe was non-significant ($.017 \pm .009$ s, $t(9) = 12.6$, $p = 0.11$).

To further assess the effect of time lag to the more recent probe on RT, we calculated the slopes of each of the lines in Figure 3a separately for each participant and performed a Bayesian t-test (Rouder, Speckman, Sun, Morey, & Iverson, 2009) on the slopes. This analysis showed "Decisive" evidence (Wetzels & Wagenmakers, 2012; Kass & Raftery, 1995; Jeffreys, 1998) favoring the hypothesis that the slopes

are different from zero (JZS Bayes Factor $= 557.9$). We did an analogous analysis to assess the effect of item lag to the more recent probe on RT. We calculated the slopes of each of the lines in Figure 3b separately for each participant and performed a Bayesian t-test on the slopes. This analysis showed "Barely worth mentioning" evidence (Wetzels & Wagenmakers, 2012; Kass & Raftery, 1995; Jeffreys, 1998) favoring the hypothesis that the slopes are not different from zero (JZS Bayes Factor $= 1.7$).

Correct RT varied sub-linearly with item lag and time lag to the more recent probe

Previous studies suggested that in relative JOR, RT varies sub-linearly with the item lag. In the context of serial exhaustive search models, this was consistent with the hypothesis that participants scan along a log-compressed representation. Due to item lag and time lag being confounded, previous studies were not able to test if the sub-linear relationship holds for both item lag and time lag. Figure 4 shows RT as a function of log time lag and log item lag. Solid line shows a logarithmic fit.

Statistical analysis confirmed that for both log item lag and log time lag, logarithmic fit is better than linear fit. Specifically, we compared a regression model of median RT onto lag to a regression model onto the base-2 logarithm of the absolute value of lag. In the time lag case, the model using $\log |\text{lag}|$ fit better than the model using $|\text{lag}|$, $\Delta LL = 5.1$, implying that the model using the logarithm is more than 150 times more likely. In the item lag case, the model using $\log |\text{lag}|$ fit better than the model using $|\text{lag}|$, $\Delta LL = 8.5$, implying that the model using the logarithm is more than 6000 times more likely. To quantify the relationship between correct RT and $\log |\text{lag}|$, we performed a linear mixed effects
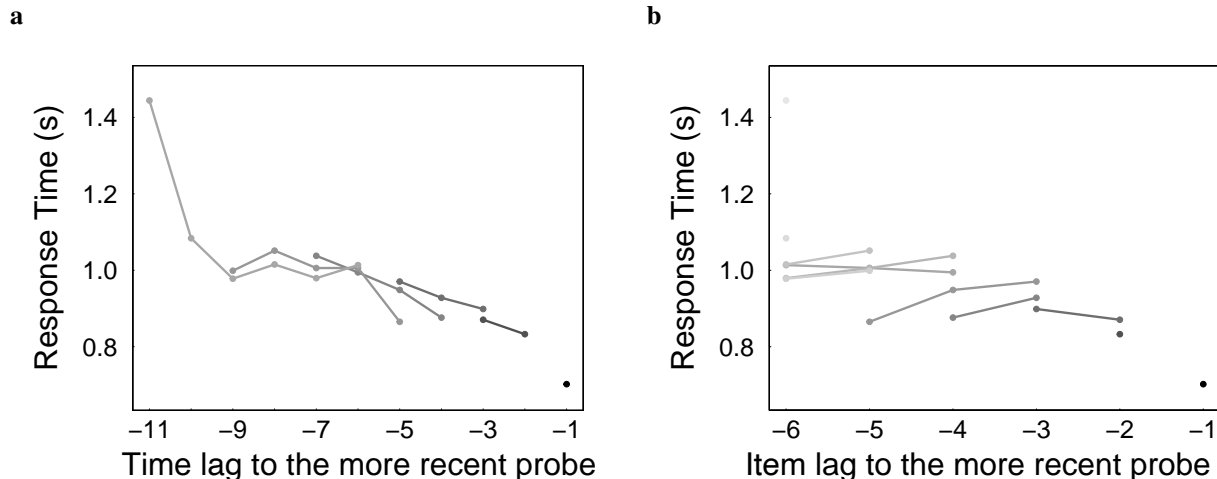
Figure 3: **RT was better explained by time to the more recent probe than by lag to the more recent probe** **a.** Shades of gray represent item lag to the more recent probe such that the darkest shade corresponds to the lag of -1 and the lightest to the lag of -6. Time on the x-axis is in units of seconds with a single time lag corresponding to the item presentation time of 0.181 s. Therefore the entire x-axis corresponds to 1.81 s. The slope of the lines illustrates the impact of time lag on RT. The significance of the slope was confirmed in subsequent statistical analysis. **b.** Shades of gray represent time lag to the more recent probe such that the darkest shade corresponds to the lag of -1 and the lightest to the lag of -11. Lines appear flat, suggesting that when time lag is taken into account, item lag does not have a significant impact on RT. This observation was also consistent with the subsequent statistical analysis.

analysis allowing for independent intercepts for each participant. In time lag case, this analysis showed that every doubling of |lag| increased RT by $.11 \pm .01$ s, $t(230) = 14.14$, $p < 0.001$. In item lag case, this analysis showed that every doubling of |lag| increased RT by $.12 \pm .01$ s, $t(143) = 13.99$, $p < 0.001$.

Discussion

In this study, we found that RT in relative JOR task was better explained by time lag than by item lag. This is consistent with (Hintzman, 2004) who came to the same conclusion in absolute JOR, using much longer lists.

Neuroscience studies have attempted to identify neural substrates for a scannable memory representation. Scannable cognitive maps spanned with bell-shaped receptive fields, such as place cells or time cells, could provide a mechanism for serial search (Behrens et al., 2018; Nieh et al., 2021). In particular, the discovery of time cells, neurons that activate sequentially following a presentation of some salient stimulus (Pastalkova, Itskov, Amarasingham, & Buzsaki, 2008; MacDonald, Lepage, Eden, & Eichenbaum, 2010; Tiganj, Cromer, Roy, Miller, & Howard, 2017) accelerated the development of neural-level models of scannable memory representations (Liu, Tiganj, Hasselmo, & Howard, 2019; Singh, Tiganj, & Howard, 2018; Tiganj, Cruzado, & Howard, 2019).

While time cells naturally support time lag as the most influential variable in accounting for RT in JOR, gating of time cells by changes in the input can give rise to a representation where RT would depend on item lag (Howard, 2014). In other words, if the sequential activation of time cells were

paused by a gating mechanism during the inter-stimulus interval, item-lag would become the most influential variable in explaining RT. Our results suggest that such gating does not happen in relative JOR task.

While we focused on discussing models of self-terminating backward scanning, since those are commonly used to explain the results in relative JOR, it is important to note that the results presented in this study do not aim to distinguish between backward scanning and other possible accounts for JOR, such as strength models, e.g., (Hinrichs, 1970; Donkin & Nosofsky, 2012). Strength models are certainly consistent with the result that RT in JOR is primarily impacted by the time lag.

In our study, lists had at most one consecutive gap between letters. This helped to balance the temporal duration of the lists, but it resulted in some missing data points for the statistical analysis. In particular, lines in Figure 3 would be longer if more than one consecutive gap was allowed. This would further strengthen the statistical results since slopes could be estimated using more data points. Future studies could attempt to address this issue by sampling a wider range of gaps.

We kept the presentation rate fixed at 5.5 letters per second. Future studies could explore whether the effects observed here are consistent if the presentation rate changes across lists. Hacker (1980) performed relative JOR at several presentation rates (however, the presentation rates were kept fixed within lists, unlike here) and observed that the overall qualitative pattern of results was robust to changes in the presentation rate. If this holds for the results described here, it would further strengthen our understanding of how the brain
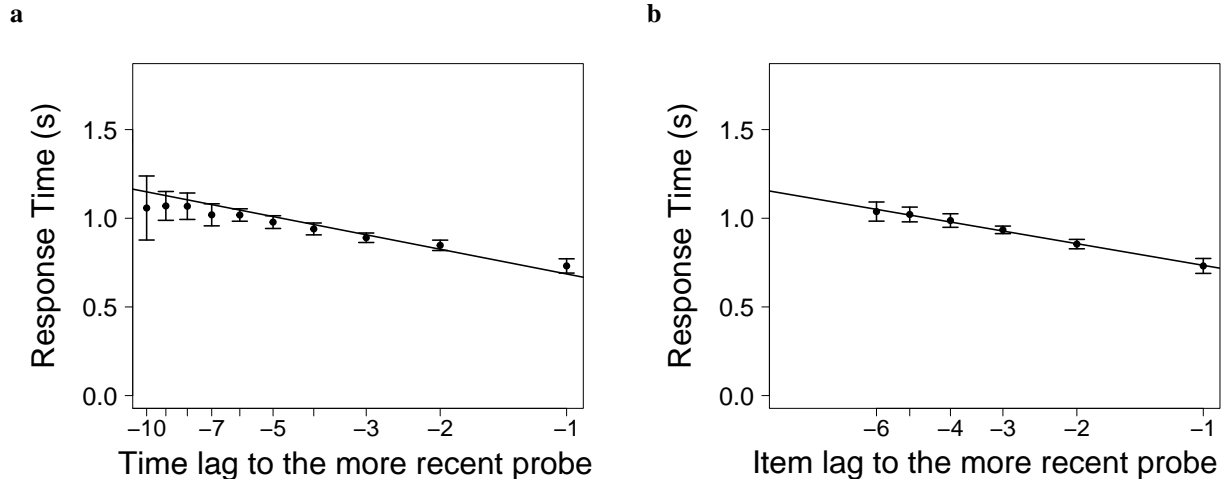
a

b

Figure 4: **Median RT varied sub-linearly with both time lag to the more recent probe (a.) and item lag to the more recent probe (b.).** Error bars represent the 95% confidence interval of the mean across participants normalized using the method described in (Morey, 2008). Solid line represents a linear fit of the data points. In plot **a.** Time on the x-axis is in units of seconds with a single time lag corresponding to item presentation time of 0.181 s.

might maintain a mental timeline of the past.

References

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.

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? organizing knowledge for flexible behavior. *Neuron*, *100*(2), 490-509. doi: 10.1016/j.neuron.2018.10.002

De Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior research methods*, *47*(1), 1–12.

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

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.

Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychological review*, *77*(3), 223.

Hintzman, D. L. (2004). Time versus items in judgment of recency. *Memory & Cognition*, *32*(8), 1298-304.

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.

Howard, M. W. (2014). Mathematical learning theory through time. *Journal of Mathematical Psychology*, *59*, 18–29.

Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773-795.

Lewandowsky, S., Brown, G. D., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language*, *54*(1), 20–38.

Lewandowsky, S., Duncan, M., & Brown, G. D. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic bulletin & review*, *11*(5), 771–790.

Lewandowsky, S., Oberauer, K., & Brown, G. D. (2009). No temporal decay in verbal short-term memory. *Trends in cognitive sciences*, *13*(3), 120–126.

Liu, Y., Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2019). A neural microcircuit model for a scalable scale-invariant representation of time. *Hippocampus*, *29*(3), 260–274.

MacDonald, C. J., Lepage, K., Eden, U., & Eichenbaum, H. (2010). Hippocampal neurons encode the temporal organization of non-spatial event sequences. *Society for Neuroscience Abstracts*, *36*, 100.15.

McElree, B., & Dosher, B. A. (1993). Serial recovery processes in the recovery of order information. *Journal of Experimental Psychology: General*, *122*, 291-315.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61–64.

Muter, P. (1979). Response latencies in discriminations of recency. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 160-169.

Nieh, E. H., Schottdorf, M., Freeman, N. W., Low, R. J.,

Lewallen, S., Koay, S. A., ... Tank, D. W. (2021). Geometry of abstract learned knowledge in the hippocampus. *Nature*, 1–5.

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*(5894), 1322-7.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237.

Singh, I., Tiganj, Z., & Howard, M. W. (2018). Is working memory stored along a logarithmic timeline? converging evidence from neuroscience, behavior and models. *Neurobiology of learning and memory*, *153*, 104–110.

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., Cruzado, N., & Howard, M. W. (2019). Towards a neural-level cognitive architecture: modeling behavior in working memory tasks with neurons. In *Cogsci* (pp. 1118–1123).

Tiganj, Z., Singh, I., Esfahani, Z. G., & Howard, M. W. (in press). Scanning a compressed ordered representation of the future. *Journal of Experimental Psychology: General*.

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064.