A distributed representation of internal time

Marc W. Howard Karthik H. Shankar
Center for Memory and Brain
Department of Psychological and Brain Sciences
Boston University

William R. Aue Amy H. Criss
Department of Psychology
Syracuse University

In press *Psych Review*, August 1, 2014, please do not quote without permission

Abstract

This paper pursues the hypothesis that a scale-invariant representation of history could support performance in a variety of learning and memory tasks. This representation maintains a conjunctive representation of what happened when that grows continuously less accurate for events further and further in the past. Simple behavioral models using a few operations, including scanning, matching and a "jump back in time" that recovers previous states of the history, describe a range of behavioral phenomena. These behavioral applications include canonical results from the judgement of recency task over short and long scales, the recency and contiguity effect across scales in episodic recall, and temporal mapping phenomena in conditioning. These tasks are believed to rely on a range of memory systems. A growing body of neural data suggests that neural representations in several brain regions have qualitative properties predicted by the representation of temporal history. Taken together, these results suggest that a scale-invariant representation of temporal history may serve as a cornerstone of a physical model of cognition in learning and memory.

The psychological hypothesis that memory and time perception are intimately related to one another has a long and rich history. For instance, Aristotle wrote "only those animals which perceive time remember, and the organ whereby they perceive time is also

The authors gratefully acknowledge support from AFOSR award FA9550-12-1-0369, NSF award BCS-1058937 and NSF CAREER award 0951612. We thank Chris MacDonald, Howard Eichenbaum, Avital Adler, Hagai Bergman, Ji Eun Kim and Min Whang Jung for sharing the data in Figure 13 and to Zoran Tiganj for sharing the results of unpublished analyses. We gratefully acknowledge helpful conversations with Bill Hoyer, Fabio Idrobo, Ralph Miller, Aude Oliva, and Roger Ratcliff. Address correspondence to Marc Howard, MarcWHoward777@gmail.com.

that whereby they remember." More recently, Tulving (1983, 1985) proposed that episodic memory results from recovery of a prior state of the time-line so that it seems as if we are re-experiencing that prior moment (e.g., Tulving, 1985). Influential quantitative models have pursued similarities between human memory for lists of words and temporal judgements (e.g., Brown, Preece, & Hulme, 2000; Brown, Neath, & Chater, 2007; Brown, Vousden, & McCormack, 2009). Similarly, a growing body of theoretical work holds that animal learning and conditioning must be understood as the learning of temporal contingencies between stimuli, leading to a deep connection between conditioning and temporal judgements (e.g., Gallistel & Gibbon, 2000; Balsam & Gallistel, 2009; Matzel, Held, & Miller, 1988; Savastano & Miller, 1998). The effort to construct a unified quantitative account of time and memory across these different disciplines has been hindered by the lack of a consensus view for how to construct a representation of internal time.

We argue that the brain ought to maintain a conjunctive representation of what happened when. This representation places events—a "what"—on an internal timeline—a "when". This concept may be more clear by means of analogy with sensory representations. For instance, in the retina different photoreceptors respond to light landing on different regions of the retina. If we consider a one-dimensional strip across the retina, the activity of the photorecptors gives illumination as a function of retinal position. Two spots of light are not represented as the average location of the two, but as two "bumps" of activity across the photoreceptors.

Sensory representations make use of an ordered representation. In the case of retinal position, the photoreceptors themselves are associated with a location. We know that one position is to the left of another, or that two receptors represent positions that are closer or farther from one another. This ordered representation is used throughout the visual system to construct a representation of what and where. Locations are accessible independently of the content they contain. For instance, we can direct attention to one particular region of retinal space in the absence of eye movements (e.g., Posner, Snyder, & Davidson, 1980) resulting in enhanced discriminability for the stimuli present in that location. Similarly, under normal circumstances, observing a particular object—a what—is accompanied by information about its location—a where.

Much like (one-dimensional) retinal position, physical time is also an ordered physical dimension. It makes sense to say that one moment is more recent than another or that two pairs of times are closer or farther from one another. We argue that our internal representation of the past respects this ordered relationship and exploits it by placing events on an internal timeline analogous to retinal position. This representation should allow attention to be focused on a subset of temporal "locations," providing enhanced access to the stimuli encoded in that part of the timeline.

Unlike sensory representations, it is non-trivial to construct a temporal history. Unlike photoeceptors located at different physical positions of the retina that respond to light in the environment, there are no "time receptors" that directly respond to stimuli from different points in the past. Some computation must intervene to allow a past state of the world to affect the current state of the memory representation. Unfortunately, obvious computational solutions have serious computational flaws that limit their usefulness in constructing models of cognition. For instance, consider a memory buffer (e.g., Atkinson & Shiffrin, 1968). As each stimulus enters, it pushes the contents of each slot in the buffer one step towards

the past. The slot number of a stimulus carries information about how far in the past it was experienced; slot number functions like a timeline relative to the present. This kind of representation is referred to as a shift register in engineering fields. While the shift register can maintain a timeline, it scales up poorly, making it ill-suited as a cognitive model of many tasks. A broad range of evidence from laboratory studies of episodic memory (e.g., Glenberg et al., 1980; Howard, Youker, & Venkatadass, 2008), interval timing (e.g., Lewis & Miall, 2009) and conditioning (e.g., Balsam & Gallistel, 2009) suggest that a representation of internal time should have similar properties over a time scales from a few hundred milliseconds to a few thousand seconds. This is not to say that memory, or timing accuracy, does not grow worse as time passes. But rather the observation is that memory, or timing accuracy, gets worse smoothly and continuously over a wide time scale rather than falling off abruptly as some scale is passed. Consider the challenge scale-invariance raises for a memory buffer. As the number of items to remember exceeds the size of the buffer, there should be an abrupt change in memory performance. The amount of resources a shift register must utilize in order to represent N time steps into the past goes up like N. As N gets large to encompass large time scales, this cost becomes probibitive. Moreover, we desire that the accuracy of the buffer decays gracefully with progressively less detailed information about the past for progressively less recent events. Getting the correct form for the decrease in accuracy is a serious challenge, even for sophisticated variants of the shift register (Goldman, 2009).

In this paper we utilize a principled computational hypothesis for how the brain could construct and maintain a scale-invariant representation of past events (Shankar & Howard, 2012, 2013). This representation provides a rich source of information that could be flexibly engaged to solve many different problems in cognitive psychology. We describe a set of such operations, including scanning, matching, and retrieval of previous states of temporal history, that could be used to query and exploit this representation to support behavior in a variety of tasks. Behavioral applications will show that this representation can be exploited to account for core findings from a variety of fields, including short-term memory, episodic memory, and conditioning. Finally, we review a growing body of evidence that suggests the brain maintains an ever-moving representation of internal time with at least some of the properties we have hypothesized.

A scale-invariant model for internal time

The mathematical details of the model for representing temporal history have been described elsewhere (Shankar & Howard, 2012, 2013). Because these details are not essential for appreciating the ability of the model to describe behavioral data we will not recapitulate them here. Rather we will emphasize the qualitative properties of the representation. The important point to appreciate here is that these properties are not simply assumed to be possible, but are the result of a well-specified computational process.

The hypothesized representation takes in an input function $\mathbf{f}(\tau)$. The function $\mathbf{f}$ depends on the particular behavioral experiment. For instance, in a list learning experiment, we might imagine that the input function is a vector-valued representation of the word presented at each moment τ. Critically, the input function only describes the currently-available stimulus; we do not assume that the memory system at time τ has direct access to previous values of $\mathbf{f}$; the goal of the memory representation is to make past values of $\mathbf{f}$

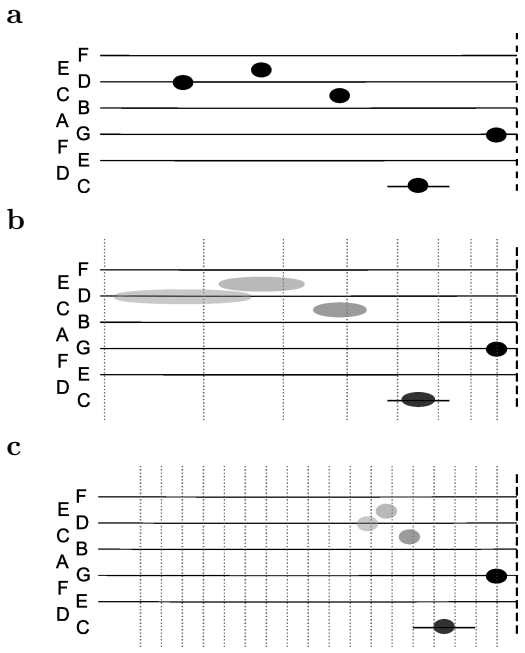

Figure 1. Schematic for a scale-invariant representation of what happened when. **a.** A series of tones, D-E-C-low C-G. unfolds in time. In musical notation, left-right position indicates the order of the notes to be played; vertical position indicates the pitch. After the sequence has been played, at a time indicated by the vertical dashed line, the history includes five tones at various points in the past. For instance, G is closer in time than the low C, which is closer than the high C. **b.** Schematic of the scale-invariant representation. Each node in a sheet corresponds to a what and a when. In this schematic figure, nodes associated with the same tone are organized in horizontal strips (corresponding to the lines and spaces of the staff); nodes associated with the same value of $\overset{*}{\tau}$, the internal time dimension are organized in vertical strips. The shading of the note schematically indicates the relative activation of the corresponding node; events further in the past result in a lower level of activation. The increasing spread for notes further in the past is intended to capture decreased temporal precision for events further in the past. Critically, the nodes are not evenly spaced along the internal time axis. Rather, if one node represents a time 1 unit in the past and the next node represents time 2, the subsequent node would represent time 4. The light vertical dashed lines show the relative spacing of nodes. **c.** Here the representation is shown schematically as a function of node number rather than internal time. Note that the light dashed lines are now evenly spaced and that the representation of events further in the past becomes compressed in internal time as well as decreasing in magnitude.

available at time τ.

Figure 1 provides a cartoon illustrating the basic idea. In the top panel, a series of tones is presented as an input pattern over time.[1] At each moment the currently sounded tone is the input. If we imagine a set of 11 nodes in $\mathbf{f}$ corresponding to the eleven notes on the lines and spaces of the staff, then the corresponding node in $\mathbf{f}$ is activated at the time each tone is played. After the entire sequence has been played (dashed line), the inputs are no longer activated, but the sequence ought to remain in memory. The goal of the representation is to estimate the history leading up to the present, enabling us to recover which tone was sounded how far in the past.

At each moment, the input function $\mathbf{f}(\tau)$ provides input to a sheet of leaky integrators $\mathbf{t}$. The sheet of nodes $\mathbf{t}$ serves as an intermediate representation. Each node in $\mathbf{t}$ receives input from a subset of the nodes in $\mathbf{f}$, allowing it to convey "what" information, and a value s that controls its time constant of integration; the fact that the set of nodes in $\mathbf{t}$ contains many different time constants can be exploited to recover "when" information from the sheet of $\mathbf{t}$ nodes.[2] Information about the history hidden in $\mathbf{t}$ is extracted by an operator $\mathbf{L}_{\mathrm{k}}^{-1}$ and written out to another set of nodes $\mathbf{T}$ at each moment:

$$\mathbf{T}(\overset{*}{\tau}) \equiv \mathbf{L}_{\mathrm{k}}^{-1}\mathbf{t}(s) \tag{1}$$

. Analogous to $\mathbf{t}$, each node in the sheet $\mathbf{T}$ corresponds to a stimulus dimension and is indexed by a value of $\overset{*}{\tau}$. The nodes in $\mathbf{T}$ are aligned with the nodes in $\mathbf{t}$: each value of $\overset{*}{\tau}$ corresponds to a value of s. The weights $\mathbf{L}_{\mathrm{k}}^{-1}$ compute the value in each $\overset{*}{\tau}$ by taking a difference of the values of $\mathbf{t}$ at neighboring values of s, closely analogous to a series of lateral inhibition circuits in a sensory system.[3] The construction of $\mathbf{L}_{\mathrm{k}}^{-1}$ depends on the value of a constant k; larger values of k result in more accurate estimates of the history.[4] We set $k = 4$ throughout this paper unless otherwise noted. For a particular value $\overset{*}{\tau}$, at any moment the set of active nodes in $\mathbf{T}(\overset{*}{\tau})$ estimates the input pattern $\mathbf{f}$ a lag $\overset{*}{\tau}$ in the past ($\overset{*}{\tau}$ is defined to be negative).

More details of the rules for updating $\mathbf{t}$, the precise specification of the weights $\mathbf{L}_{\mathrm{k}}^{-1}$, and the rationale for why this approach works are described in depth elsewhere (Shankar & Howard, 2012, 2013). Here we simply describe key properties of the representation $\mathbf{T}$ in simplified form and attempt to provide an intuition for why these properties are useful.

Nodes in the representation respond to conjunctions of what and when. The middle panel of Figure 1 shows a cartoon depiction of the history $\mathbf{T}$ at the end of the sequence. Each row (i.e., each line and space of the staff) corresponds to a set of nodes that respond to one tone. Each column corresponds to a set of nodes that correspond to a particular value of past time. More recent times are to the right and less recent times to the left; the spacing of this cartoon is chosen such that the values of $\overset{*}{\tau}$ are lined up with the value of physical

[1] This motif may be familiar to movie fans.

[2] Each node in $\mathbf{t}$ is activated by its input and then decays exponentially with a time constant controlled by its value of s. At time τ, the set of all nodes in $\mathbf{t}$ encodes the Laplace transform of the history leading up to the present, $\mathbf{f}(\tau' < \tau)$.

[3] More precisely, $\mathbf{L}_{\mathrm{k}}^{-1}$ estimates the kth derivative with respect to s.

[4] Post (1930) proved that in the limit as k goes to infinity, $\mathbf{L}_{\mathrm{k}}^{-1}$ implements the inverse Laplace transform. For finite k, the inversion is imprecise, resulting in a smear in the estimate of the history.

time to which they correspond. Taking the pattern of active nodes across a column at any one moment gives an estimate of the stimulus that took place at that time in the past. Taking a row across $\mathbf{T}$ gives an estimate of the history of that particular tone as a function of past time. Peaks appear at values of $\overset{*}{\tau}$ corresponding to remembered occurrences of that tone.

Notice that the pattern of active nodes across the sheet changes as time unfolds. Consider a particular node coding for a particular tone a particular $\overset{*}{\tau}$ in the past. That node will not become activated unless the history contains its preferred stimulus at the appropriate time in the past. This means that after the tone is presented, the node will not respond immediately, but after some characteristic delay (determined by the node's value of $\overset{*}{\tau}$). Because there are many nodes that respond to any particular tone with a variety of different values of $\overset{*}{\tau}$, presenting the tone sets off a sequence of nodes in $\mathbf{T}$ as they become activated at different delays. This sequence ought to repeat if the same tone is repeated at a later time.

The representation of history is less accurate for events further in the past. The representation of history represents recent events with more precision than less recent events. This is manifest in two ways. First, the peak of activity is less pronounced for events further in the past; Figure 1 shows this in cartoon form as lighter shading for events further in the past. Second, the estimate of the internal time at which an event took place becomes less precise; Figure 1 captures this in cartoon form as wider and overlapping patterns of activity.

These properties can be seen more precisely by looking at the actual pattern of activity across values of internal time in Figure 2. The middle panel of Figure 2 shows bumps of activity across nodes. Each line corresponds to the activity of a set of nodes coding for a particular stimulus. The most recent stimulus has caused a bump of activity that is the farthest to the right; less recent stimuli cause bumps of activity that peak progressively further to the left. The change in overall level of activity can be seen by noting that the peak of each curve becomes progressively lower for stimuli further in the past. This decrease is accompanied by a spread in the peak along the $\overset{*}{\tau}$ axis. The overlap in the curves also increases as the stimuli recede further into the past.

This spread, looked at from the perspective of a single node, means that the "temporal receptive field" for nodes with larger values of $|\overset{*}{\tau}|$ coding for times further in the past, is broader than the temporal receptive field for nodes with smaller values of $|\overset{*}{\tau}|$ coding for times nearer to the present. This means that after presentation of a single stimulus, nodes with larger values of $|\overset{*}{\tau}|$ are activated for a longer amount of time than nodes with smaller values of $|\overset{*}{\tau}|$ as the stimulus recedes into the distant past. The decrease in the peak trades off with this longer duration of firing such that the total activation summed over time is constant for all cells in $\overset{*}{\tau}$. The spread in the temporal receptive fields shows scalar variability, such that the spread goes up linearly with $|\overset{*}{\tau}|$; the peak activation goes down with $|\overset{*}{\tau}|^{-1}$.

Internal time uses Weber-Fechner spacing of nodes. Nodes in $\mathbf{T}(\overset{*}{\tau})$ representing events further in the past have wider temporal receptive fields than nodes representing more recent events. In sensory systems, an analogous systematic widening of receptive fields is commonplace. For instance, in the visual system, cells with receptive fields further from the fovea

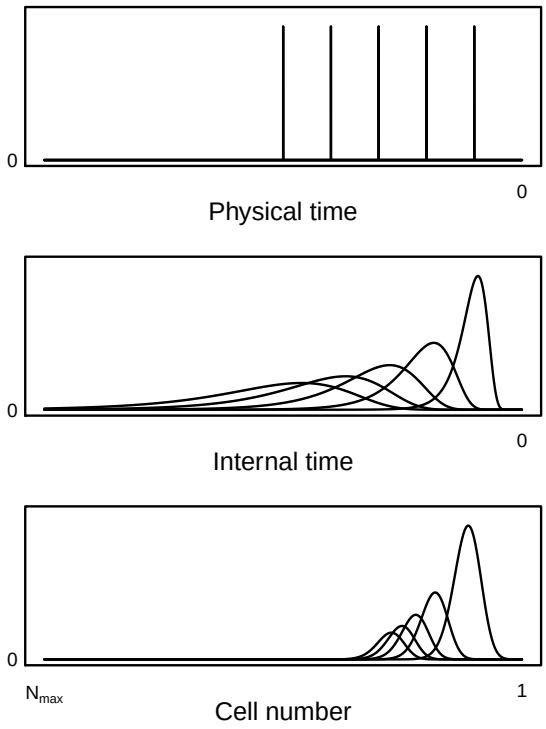

Figure 2. More detailed depiction of the scale-invariant representation of internal time. Top: Five stimuli, corresponding to the five tones in Figure 1 were presented in the past. Middle: At a time after presentation of the last tone, the activation across nodes representing internal time are shown for each of the five tones. Each line corresponds to the activation of the set of nodes corresponding to that tone (i.e., a row in Figure 1). Stimuli further in the past are represented by bumps of activity with lower peaks and broader uncertainty. Bottom: The same set of activations in the middle panel is shown as a function of node number rather than along the internal time axis. This is the same distinction as between the bottom two panels of Figure 1. Because the nodes are logarithmically spaced, the time axis is "foreshortened" relative to node number. When viewed as a function of node number, the bumps of activity still have decreasing amplitude but now have the same relative width. In this schematic figure, $k = 12$ for visual clarity.

are broader than receptive fields close to the retina. However, cells with wide receptive fields that have centers close together are redundant with one another. It is adaptive to have a cell density that is inversely related to receptive field width (and vice-versa). Note that indeed the density of photoreceptors on the retina decreases with distance from the fovea. We can describe the "location" of a given photoreceptor along a one-dimensional strip of the retinal surface either in terms of its physical position along the retina, or in terms of its relative cell number, starting at 1 at the center of the fovea and incrementing the count every time we encounter another cell as we move out radially. These two numbers are in one-to-one correspondence. However, because the density of receptors is not constant, the relationship between physical location and cell number is not linear.

We propose that the spacing of nodes supporting the $\overset{*}{\tau}$ axis is not constant, but instead has Weber-Fechner law scaling, such that the Nth set of nodes is associated with a $\overset{*}{\tau}$ that goes like $N = A \log |\overset{*}{\tau}| + B$, where A and B are some constants.[5] This relationship implies that the spacing between adjacent nodes goes up proportional to $\overset{*}{\tau}$. The middle panel of Figure 1 shows the relative spacing of cells (thin vertical lines) on the $\overset{*}{\tau}$ axis; the bottom panel shows, in cartoon form, how the particular history in the middle panel would look in terms of node number. The argument that Weber-Fechner scaling is optimal can be made in several ways (for detailed information theoretic arguments see Shankar & Howard, 2013). However, many of these arguments hinge on a simple observation—that the shape of the bump of activity in $\mathbf{T}$ as a function of node number does not change shape as the stimulus recedes into the past (Fig. 2, bottom). This means that the discriminability of adjacent nodes is constant for all values of $\overset{*}{\tau}$, equating redundancy across the set of nodes. As a consequence of Weber-Fechner law spacing, the longest time scale that can be represented goes up exponentially as a function of the number of nodes dedicated to the representation.[6] This is a tremendous savings in terms of resources relative to a shift register where the number of nodes would go up linearly with the longest time scale that can be represented.

The representation of history is scale-invariant. Natural signals contain potentially useful information over a wide range of temporal scales (e.g., Alvarez-Lacalle, Dorow, Eckmann, & Moses, 2006; Sreekumar, Dennis, Doxas, Zhuang, & Belkin, 2014; Voss & Clarke, 1975). Mathematical scale-invariance, the property that memory does not have a characteristic scale, has been argued to be a central principle of cognitive psychology (Anderson & Schooler, 1991; Chater & Brown, 2008; Kello et al., 2010) and is a key feature of many behavioral models of timing, conditioning and episodic memory (Brown et al., 2000, 2007; Gallistel & Gibbon, 2000; Gibbon, 1977; Miall, 1989). It can be shown that the representation of history we have just described is mathematically scale-invariant (Shankar &

[5] These constants are not identifiable from behavioral data and will not be considered further.

[6] It is a non-trivial computational challenge to implement a temporal dimension with Weber-Fechner law spacing. For instance, consider a shift register in which chains of cells connect one to the other to cause a "bump" of activity to gradually move back in time (e.g., Goldman, 2009). In order to implement Weber-Fechner law spacing, the rate at which one node communicates with its neighbors must change dramatically as a function of cell number. Even if this could be accomplished, these "chains" are still sensitive to disruption and noise. Because $\mathbf{T}(\overset{*}{\tau})$ is constructed directly from $\mathbf{t}$ in the neighborhood of a particular value of s, there is not a chain of information that flows from one value of $\overset{*}{\tau}$ to another. Choosing the spacing of $\overset{*}{\tau}$ nodes amounts to choosing the spacing of the values of s.

Howard, 2012). The spread in the temporal representation is proportional to the time in the past that is being represented. This property is computationally non-trivial[7] and is consequence of the specific computational mechanism used to construct $\mathbf{T}$. By constructing different behavioral applications from this scale-invariant representation we can account for scale-invariant behavior in some circumstances but scaled behavior in other circumstances. The converse—constructing a scale-invariant behavioral model from a scaled memory representation—is very difficult; scale-invariant behavior all-but-requires a scale-invariant memory representation.

Accessing and exploiting internal time to support behavior

We propose that subjects maintain at each moment an ordered representation of internal history that records the stimuli experienced in the past. However, this representation is not necessarily queried at each moment, nor is the mode of access necessarily automatic. Rather, the representation is queried and exploited using a variety of methods appropriate to the demands of the cognitive task at hand. Consider the analogy to the ordered representation of retinotopic position available to the visual system. While patterns of light might be available at all parts of the retina at a particular moment, this does not require that we direct our attention uniformly to the entire retinal display. Based on our goals at a particular moment we may choose to direct attention to a particular region, giving us enhanced access to the information present in that part of the retina. Alternatively, we may choose to retrieve from memory (or even imagination) a previously-experienced image which is then in some sense projected onto the "mind's eye," retrieving the ordered relationships between the components of the recovered image. Also note that these different operations may rely on very different brain regions. How we choose to utilize the visual information available at any moment depends on our goals, but the essential point is that these operations are only possible because of the existence of an ordered representation of retinal position.

We hypothesize that an analogous set of operations on an ordered representation of temporal history could be used to generate performance in a broad variety of memory tasks. We construct a variety of behavioral models that draw on the representation of internal time in task-appropriate ways. We restrict our attention to three basic operations that could be utilized to support behavioral performance: scanning, matching, and a "jump back in time." Scanning is analogous to the direction of visual attention to a sequence of retinal locations, sweeping from the present towards the past. Matching compares the overlap, summed over past times, between two states of temporal history. A "jump back in time" causes a past state of temporal history to be recovered for use in the present. This recovered state could then be used to match (or, in principle at least) scan as appropriate to drive an appropriate response.

Scanning: Sequential examination of internal past time. Many authors have made the analogy between time and a perceptual dimension (e.g., Brown et al., 2007; Crowder, 1976; Murdock, 1960, 1974). It is clear that subjects can choose to strategically direct

[7] Consider the result of averaging noise in a chain. Under a wide range of circumstances, we would expect the variability to be subject to the law of large numbers. The law of large numbers implies that the standard deviation of the distribution should go with the square root of the number of links in the chain, rather than linearly with the number of links in the chain.

focused attention to a region of a particular perceptual dimension—for instance a region of retinotopic space or a band of frequencies. Taking the analogy between ordered perceptual dimensions and the ordered representation of history seriously, we propose that subjects can choose to access contiguous subsets of the ordered representation. Summing $\mathbf{T}(\overset{*}{\tau})$ over a region on the ordered "when" dimension results in a vector across stimulus dimensions—i.e., a "what" vector in the same space as $\mathbf{f}$. Critically, we assume that the summation across the temporal dimension is with respect to Weber-Fechner-spaced node number (Figure 1c, 2 bottom) and not with respect to the internal time axis (Figure 1b, 2 middle). The "what" output of this operation can then be compared to a probe stimulus and used to drive a decision process.

This framework is sufficiently broad to encompass serial scanning models and parallel access models. Serial scanning is analogous to sliding this window of "attention" smoothly, one set of nodes at a time. Parallel access is analogous to generating a more broad window of attention over a range of contiguous nodes in $\mathbf{T}$. We will utilize sequential scanning in a behavioral model of the judgment of recency (JOR) task below. Because of the Weber-Fechner spacing of the internal time axis, the number of nodes that need to be traversed to reach a particular time $\overset{*}{\tau}$ in the past should go up like $\log(\overset{*}{\tau})$ (Hinrichs, 1970). If one were to access a broader region of past time in parallel, the strength of each stimulus representation in $\mathbf{T}$ falls off like a power law function of the time since it was presented (e.g., Donkin & Nosofsky, 2012). This last result follows from the Weber-Fechner spacing of the nodes supporting the temporal axis. The area under the curves in the bottom panel of Figure 2 give an intuition into why this property holds. Because each curve has precisely the same shape as a function of node number, and because the peak values are scaled by τ_o^{-1}, the area under the curve also falls off like a power law.

Matching states of temporal history. In some circumstances it is advantageous to base a decision or action not directly on the contents of the current $\mathbf{T}$, but on how well the current $\mathbf{T}$ matches some previously stored state of $\mathbf{T}$. Consider the case in which a stimulus A is presented and ten seconds later B is presented; A and B could be the CS and US in a Pavlovian trace conditioning experiment. The state of history at the moment when B is presented, lets call it $\mathbf{T}_{\text{B}}$, includes A ten seconds in the past. Now suppose A is provided as a cue long after the original presentation of the pairing of A and B. It would be advantageous to be able to predict that B will occur at the correct time. But note that the current state of temporal history after A is presented as a cue contains no information about B *per se*. However, suppose an associative memory was constructed whereby the state of history when B was originally presented, $\mathbf{T}_{\text{B}}$, is associated to B.[8] This association means that a state of history predicts B to the extent it matches B's encoding history, $\mathbf{T}_{\text{B}}$. After A is repeated, A enters the current state of the history and begins receding into the past. Shortly after A is repeated, the match between the current state of $\mathbf{T}$ and $\mathbf{T}_{\text{B}}$ is weak; $\mathbf{T}_{\text{B}}$ includes A ten seconds in the past, not one or two seconds in the past. As A recedes further into the past, the match to $\mathbf{T}_{\text{B}}$ increases as A approaches ten seconds in the past. In this way, the match between a stored state of history and the current state of history can support an appropriately-timed behavioral association between A and B.

[8] In Shankar and Howard (2012) this was accomplished *via* a simple Hebbian outer product association between $\mathbf{T}$ and $\mathbf{f}$, but other approaches are possible.

Because of the "blur" in the time-line, the match between a state and a subsequent state changes gradually from moment to moment. We will use the match between stored states of $\mathbf{T}$ and the current state of $\mathbf{T}$ to drive performance in free recall and conditioning tasks.

Jump back in time. The third operation we make use of is the recovery of a previous state of temporal history. Suppose that A is presented at a specific point in time, preceded by some stimuli. Let us refer to the state of history immediately preceding the presentation of A as $\mathbf{T}_{\text{A}}$. We propose that if A is repeated at a later occasion, then under some circumstances this can cause the partial reconstruction of $\mathbf{T}_{\text{A}}$, making information about the stimuli preceding the initial presentation of A available in the present. The recovery of a previous state of history is analogous to the "jump back in time" hypothesized to support episodic memory (Tulving, 1983). The recovery of a previous state of temporal context is a prominent feature of several detailed models of episodic recall (Sederberg, Howard, & Kahana, 2008; Polyn, Norman, & Kahana, 2009; Howard, Jing, Rao, Provyn, & Datey, 2009), where this mechanism enables a behavioral account of backward associations in episodic recall (e.g., Kahana, 1996; Kiliç, Criss, & Howard, 2013) as well as the ability to integrate fragmented episodes into an integrated representation (Howard, Fotedar, Datey, & Hasselmo, 2005; Howard, Jing, et al., 2009). For instance, consider a double function list where subjects study a pair A B and then later study a pair with an overlapping stimulus B C. Subjects show a transitive association between A and C even though those stimuli were never presented close together in time (Bunsey & Eichenbaum, 1996; Primoff, 1938; Slamecka, 1976). This finding makes sense if the second presentation of B causes a "jump back in time" recovering the temporal context in which it was previously experienced; the temporal context of B includes information about A, providing a mechanism for A to become associated to C.

In the behavioral applications we use here, recovery of a scale-invariant temporal history affords several advantages over the recovery of a temporal context vector as used in the temporal context model (TCM). First, because the representation of history is scale-invariant and the temporal context vector is not (Howard, 2004), recovery of the temporal history can result in behavioral contiguity effects that persist over long periods of time. Second, because the state of temporal history, contains a representation of what happened when[9] integrating different histories enables the formation of a temporal map that provides information about temporal relationships between different events (Cole, Barnet, & Miller, 1995). That is, rather than simply being able to learn that A "goes with" C, information about the temporal relationships between stimuli can be extracted and integrated. As with the other operations for accessing and exploiting internal time, we have not provided a detailed mechanism by which a jump back in time could be accomplished. We simply assume that the recovery is accomplished somehow at some appropriate times. We emphasize that this "jump back in time" is not an obligatory consequence of repeating a stimulus.

[9] In TCM, the temporal context vector is a leaky integration of the inputs caused by the items presented. As such, the history of a single item is described by a single scalar value rather than a function of prior time. The temporal context vector is thus an ahistorical strength.

Overview of behavioral applications

We start by building a simple behavioral model using sequential scanning of $\mathbf{T}(\overset{*}{\tau})$ and apply this behavioral model to JORs over short (a few seconds, e.g., Hacker, 1980) and long (several dozen seconds, e.g., Yntema & Trask, 1963) laboratory time scales. In the next subsection we build a simple behavioral model of recency and contiguity effects in episodic recall across a range of time scales. This model makes use of matching and the recovery of prior states of temporal history. Finally, we use matching and the recovery of previous states of temporal history to account for temporal mapping phenomena observed in second order conditioning experiments (e.g., Cole et al., 1995).

Judgments of recency across time scales

The JOR task taps subjects' estimates of the time in the past at which an event took place. In a relative JOR task, participants are presented two probe items and asked to select the probe that occurred more recently. In an absolute JOR task subjects are presented a single probe and judge the time, or number of stimulus presentations, that have passed since its most recent presentation. In many short-term JOR experiments (Hacker, 1980; Muter, 1979; Hockley, 1984; McElree & Dosher, 1993), subjects perform relative JOR on a list of common stimuli (e.g., letters) presented rapidly one at a time over a few seconds. The pattern of response times (RTs) from short-term JOR have been taken as strong evidence for scanning. In the short-term JOR task, when subjects correctly select the more recent probe item, their RT is a monotonic function of the recency of the more recent probe such that more distant items had longer RTs. However, there is little or no effect of the recency of the less-recent probe item on correct RT (Hacker, 1980; Hockley, 1984; McElree & Dosher, 1993; Muter, 1979). In contrast when subjects incorrectly choose the less-recent probe, RT is a monotonic function of the recency of that item, but is only minimally affected by the recency of the more recent probe. The top panel of Figure 3 shows representative data from the Hacker (1980) study. These results, and several other related findings, suggest that subjects examine their memory by starting at the present moment and serially scan to earlier times to find the probe stimuli, stopping when one of the probes is matched by the contents of memory (Hacker, 1980; Hockley, 1984; McElree & Dosher, 1993; Muter, 1979).[10]

Here we develop a simple behavioral scanning model based on sequential access of the representation of internal past time. This simple model describes short-term JOR data from the Hacker (1980) paradigm over a few seconds. Because the representation is scale-invariant, there is no reason the same behavioral model couldn't also extend to longer laboratory time scales. We show that the same simple behavioral model also accounts for key features of JORs observed over time scales up to a few minutes (Hintzman, 2010; Yntema & Trask, 1963).

The scanning model is very simple. Each point in the time-line $\mathbf{T}(\overset{*}{\tau})$ is sequentially accessed starting from the present and moving subsequently to progressively less recent

[10]The scanning interpretation is not only supported by mean RT. Hockley (1984) found that the entire RT distribution shifted as the more recent probe was moved further into the past (see also Hacker, 1980; Muter, 1979). Similarly, McElree and Dosher (1993), using the response signal procedure, found that the time at which information began to accumulate depended on the recency of the probe.

values of $\overset{*}{\tau}$. With reference to the bottom panel of Figure 2, we start at the right hand side of the figure. At each step we query the representation to determine what stimulus was presented at that point in time. At the next time step we move one node to the left and repeat the query. The output of the contents of memory is compared to the probe stimulus in absolute JOR, or the two probe stimuli in relative JOR. At each step of the scan, the decision process terminates with probability a times the match between $\mathbf{T}(\overset{*}{\tau})$ and the corresponding probe. a is a free parameter. In a relative JOR task, the subject chooses the probe whose process terminates first. If the search goes on long enough with neither probe being selected, the subject guesses randomly (Hacker, 1980). In absolute JOR, the model is the same except there is only one decision process corresponding to the one probe stimulus. In absolute JOR we assume that the subject reports the node number on which the search terminates and uses that information as the basis for a response. Recall that node number goes up logarithmically with $|\overset{*}{\tau}|$ (Fig. 2, bottom). Appendix A provides a more formal description of the scanning model and several useful results.

A few comments are in order to give the reader a stronger intuition before comparing the model to data. First, the smear in the representation $\mathbf{T}(\overset{*}{\tau})$ makes it easier to confuse pairs of events further in the past than recent events. That is, the representation of two events separated in time by a fixed duration overlap less if the two events are close to the present and overlap more as the pair of events recedes into the past. Because the probability of termination depends on the value of $\mathbf{T}(\overset{*}{\tau})$, more overlap means that there is a greater probability that the incorrect probe will be chosen. Second, more recent stimuli are more likely to return information even if there is only one probe competing. This is because the peak activation caused by a studied item is a decreasing function of the delay since it was presented. Both of these factors make memory worse for stimuli further in the past. Third, the Weber-Fechner spacing of the $\overset{*}{\tau}$ axis causes the scanning model to show a logarithmic increase in scanning time as a probe becomes less recent. Moreover, the scanning model of absolute JOR generates logarithmic increase in rated recency as a function of actual recency (Hinrichs & Buschke, 1968).

Serial scanning in short-term memory. The bottom panels in Figure 3 show predictions of the simple scanning model described above for the Hacker (1980) results. The parameter a was manipulated to provide the best fit to the experimentally-observed accuracy. We calculated the probability of a correct response by numerically integrating the appropriate expressions (see Appendix A). To estimate RT, we calculated the expectation of the number of the node at which the search terminated. There were no parameters manipulated to try and fit the RT data *per se*, but several choices that had to be made to generate reasonable numbers. We set the guess RT to the log of three times the $\overset{*}{\tau}$ corresponding to the longest list length to roughly correspond to the parameters of the experiment. We also added a constant to the expectation of $\log|\overset{*}{\tau}|$ to represent the lower limit of integration (i.e., the value of $\overset{*}{\tau}$ closest to zero) and the non-decision time. This affected all RTs by the same constant amount and had no effect on their relative spacing.

Despite not being fit to RT, the scanning model captures the basic pattern of results for RT for both correct responses and errors. Correct RTs depend a great deal on the lag of the more recent probe and only minimally on the lag of the less recent probe. In

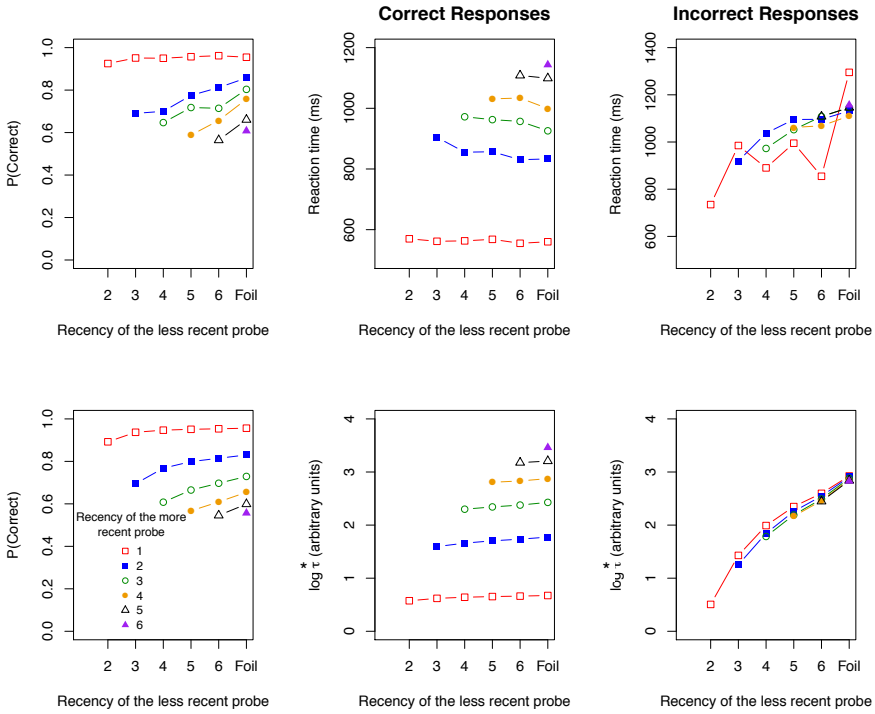

Figure 3. Short-term memory scanning. In Experiment 2 of Hacker (1980), participants were shown a list of consonants at a rate of one item every 180 ms. At the end of the list, they were shown two probe stimuli and asked to select the most recent item. The top row shows experimental results for accuracy (top left), response time (RT) for correct responses (top center), and incorrect RTs (top right). Note especially that the recency of the less recent probe has little or no effect on correct RT, but has a dramatic effect on error RT. The bottom panels contain corresponding model fits using a simple scanning model over internal past time. Memory for each of the probes was queried by scanning successive points of the timeline $\mathbf{T}$. The probe that returned useful information first was selected. If neither probe returned useful information when $\overset{*}{\tau}$ reached three times the list length, the model guessed. RT is estimated as the number of nodes ($\log|\overset{*}{\tau}|$) that were scanned prior to the decision. The parameter controlling the instantaneous probability of terminating the search was adjusted to fit the accuracy data ($a = 2.69$).

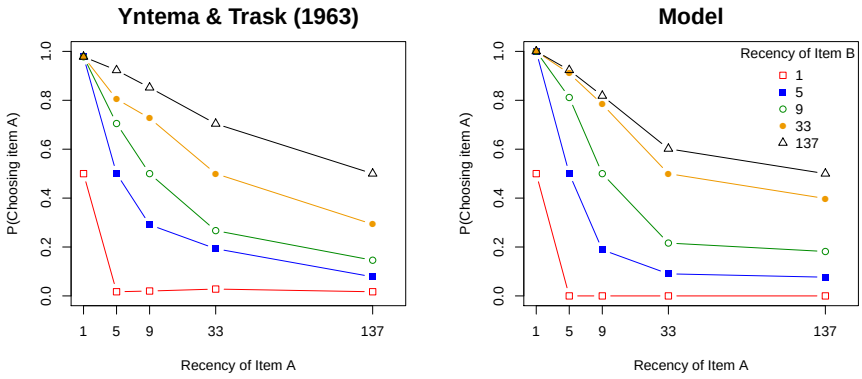

Figure 4. Long-term memory scanning. In Experiment 2 of Yntema and Trask (1963), subjects were presented with a continuous stream of pairs of words. On test trials, they were shown two probes and asked to select which one was presented more recently. The plot gives the probability of choosing item A as the more recent item (either correctly or incorrectly) as a function of its recency. The separate lines represent the recency of the other probe item, B. The model predictions are for the same scanning model as in Figure 3 with $a = 9.71$.

contrast, error RTs depend greatly on the recency of the less recent probe and almost not at all on the recency of the more recent probe.[11] The model deviates slightly from the data in that the correct RT functions are not precisely flat, trending slightly away from a distance function. The model produces this behavior for the same reason as the Hacker (1980) behavioral model—less recent items are missed at a higher rate than more recent items. As a consequence, the proportion of correct responses attributable to guesses goes up as the recency of the less recent item decreases. Because the guess RT is big relative to the non-guess RT, the result is a slight increase in RT.

Serial scanning in long-term memory. In this subsection, we show results from two experiments that illustrate important properties of the scanning model with respect to JORs over longer laboratory time scales. First, we show that the simple scanning model captures the basic pattern of accuracy as a function of lag in a relative JOR task (Yntema & Trask, 1963). Second, we show that the scanning model is able to account for the separate effect of multiple presentations of an item, a key feature of the data that lends support to multiple-trace models of the JOR task (Hintzman, 2010).

Yntema and Trask (1963) performed a continuous relative JOR task over a range of time scales up to several dozen seconds. On each study trial, a pair of words was presented. On each test trial, a pair of probe words was also presented, with the subject required to judge which of the probes was presented most recently. In a relative JOR, there are two relevant delays associated with the test of the two probes. Let us refer to the difference between the time of presentation of the two probes as the lag and the delay from the more recent probe at the time of test as the retention interval. In the relative JOR task, accuracy

[11] It should be noted that the apparent increase in the empirical RT when the probes are close together in the list did not hold up to statistical scrutiny (Hacker, 1980) and is not present in replications of this study by other authors (Hockley, 1984; Muter, 1979; McElree & Dosher, 1993).

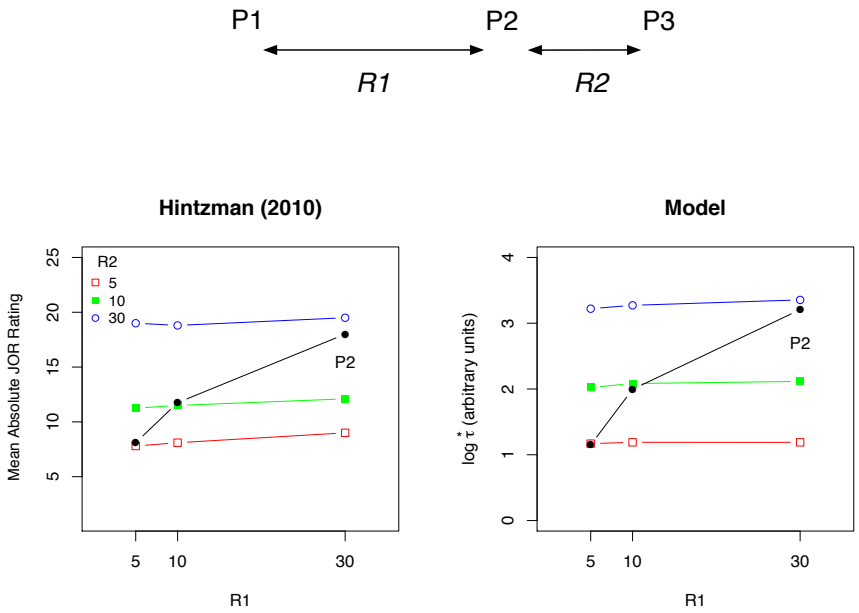

Figure 5. Relative independence of separate presentations of an item. Top: In Experiment 1 of Hintzman (2010), words were presented on three occasions, P1, P2, and P3. At each presentation, the subject had to indicate whether the word had been presented earlier and, if so, estimate its recency. The recency of the word when it was repeated at P2 is denoted R1. The recency when the word was repeated the second time at P3 is referred to as R2. R1 and R2 were set to 5, 10, or 30. The black line with filled circles gives performance on P2. The other symbols give performance on P3 (see legend). Bottom, Left: Empirical results for the absolute recency judgement for probes judged old as a function of R1. Note that R1 has little or no effect on the judgements at P3. Bottom, Right: The mean value of $\log|\overset{*}{\tau}|$ at which the scan through history returned. $a = 15$.

is relatively good when the retention interval is small compared to the lag; as the retention interval gets large relative to lag, accuracy decreases (e.g., Frey & Fozard, 1970; Lockhart, 1969; Wasserman, DeLong, & Larew, 1984; Yntema & Trask, 1963). The basic findings (Figure 4) are that, all other things being equal,

1. accuracy decreases with the retention interval to the more recent probe and
2. accuracy increases with the lag between the two probes.

The scanning model captures both of these properties quite well (Figure 4). Note that the self-terminating scanning model used here would predict latency differences in long-term relative JORs similar to those observed in the Hacker (1980) data. To our knowledge, there have been no reports of latency data in relative JORs over the time scale of dozens of seconds.

The Yntema and Trask (1963) findings are perfectly consistent with a strength account and do not independently require a temporal representation. However, the fact that subjects show memory for separate presentations of an item argues strongly against a strength account of JOR performance. Hintzman (2010) presented subjects a continuous absolute

JOR task. Words were presented three times; we refer to the three presentations as P1, P2 and P3. On P1, the first presentation of an item, subjects were to judge it as new. On P2 and P3, subjects were to rate it as old. For words the subjects correctly judged as old, they were to rate its absolute recency. Let us refer to recency of the probe the second time it is presented as R1 and the recency of the probe the third time it is presented (i.e., the difference between P2 and P3) as R2 (see Figure 5, top). At P2, subjects rated R1 for the words they judged as old. Their ratings display an approximately logarithmic judgement for the words they rated as old (Figure 5a, black line). At P3, there are two prior presentations of the probe, one at recency R2, and the other at recency R2+R1. As R1 and R2 are factorially varied, the question is whether R1 has an effect on the judgement at R2. The nearly horizontal colored lines in Figure 5a show that there is a minimal effect of R1 on the judgement of R2.

Because the Hintzman (2010) experiment only reported JOR ratings for probes that were also judged old, we only included results for searches that returned useful temporal information. The false alarm rate to the recognition judgement on the first presentation in the Hintzman (2010) data was only .03 (the hit rate on P2 was .93; on P3 it was .98) so it is reasonable that very few judgements were based on guessing. The relatively high value $a = 15$ for the parameter controlling the instantaneous probability of the search returning means that the probability of a search returning a useful value at any point during the scan is correspondingly high. We assume, but do not model, a linear mapping between N, the output of a particular search, and the subjects' understanding of physical time, leaving the answer in terms of $\log|\overset{*}{\tau}|$.[12]

Figure 5 shows that the simple behavioral model based on the scale-invariant representation of temporal history captures the basic trends in the data. It describes the form of the increase in the rating at P2 with R1. It describes the strong effect of R2 on the recency judgement at P3 and the relatively minimal effect of R1 on the judgement at P3. It even describes the slight increase in the judgement at P3 as R1 increases. In most cases, the query is already returning information due to the most recent presentation of the item. Because the search is self-terminating and there is a strong recency effect, additional information about the earlier presentation does not contribute much to the judgement. However, it does tend to move the judgement slightly towards the past.

Discussion. We showed that a one-parameter backward self-terminating scanning model built on top of the distributed representation of temporal history captures canonical effects from the JOR task. The model smoothly transitions between putative short-term (Hacker, 1980) to long-term (Yntema & Trask, 1963) results using a scale-invariant memory representation. With respect to the account of short-term JOR data, the model differs very little from previous scanning models, but the representation of internal time, with its Weber-Fechner spacing, makes predictions about longer-term JOR that are not shared by previous treatments of short-term JORs. For instance, the fits for the Hacker (1980) data of the model used here are very similar to those from the scanning model in Hacker (1980). However, in that model the availability of the stimuli were allowed to vary as free parameters—availability was found to fall off with recency. Here, that recency effect

[12]We have also added a constant, corresponding to the lower limit of integration, to avoid reporting negative values of RT.

emerges from properties of the representation of internal time and generates predictions for longer-term JORs as well. Similarly, Brown et al. (2000) constructed a behavioral scanning model of short-term JORs built from the coupled oscillators in the OSCAR framework. In OSCAR, the state of the oscillators during study is associated to each list item. The similarity of the state of the oscillators at the end of the list is used to generate a strength for each list item. The strengths are then sorted and the subject scans through the list items weighted by their strength until a match to a probe is found. This is something of a quasi-scanning model—sequential access is assumed to be taken on an ahistorical strength. While the predictions of the present model are quite similar for the Hacker (1980) data, the quasi-scanning model based on OSCAR would tend to make counterfactual predictions for the Hintzman (2010) experiment. Because both P1 and P2 would contribute to the strength, the near-independence of absolute JOR for the most recent presentation on the recency of the previous presentation (Figure 5) is difficult to reconcile with an ahistorical strength representation.[13]

Although the present behavioral model does a reasonable job of accounting for behavior in both short-term memory and long-term memory, this is not to say that the behavioral model, nor behavior across the tasks, is scale-invariant. The relationship between a scale-invariant representation and behavior is more nuanced, as it depends on the behavioral model used to generate behavior. The scanning model used here has a scale set by the parameter a, which controls the instantaneous probability of the search returning, and exhibits different properties in different regimes. When the recency of a probe item $\tau_o \ll a$, memory scans return useful information almost certainly. In this regime, different probes compete with one another. On the other hand, when $\tau_o \gg a$, memory searches do not often return information and the searches behave as if they were independent. Although the underlying representation is precisely scale-invariant, the behavioral model exhibits qualitatively different behavior in these two regimes.

Scale-invariant recency and contiguity in free recall

In the free recall task, subjects are presented with a list of words and asked to recall as many of the words as they can in the order they come to mind. Data from the recency effect in free recall (Glanzer & Cunitz, 1966; Murdock, 1962) were instrumental in the adoption of the distinction between short-term and long-term memory (Atkinson & Shiffrin, 1968; Glanzer, 1972). The immediate recency effect was taken as evidence of the enhanced availability of items that remained active in short-term memory. Inserting a delay at the end of the list (Glanzer & Cunitz, 1966; Postman & Phillips, 1965) sharply attenuates the recency effect; according to models of short-term memory, this attenuation happens because the end-of-list items are displaced from short-term memory by the distractor.[14] Subsequent evidence on the long-term recency effect (Bjork & Whitten, 1974; Glenberg et al., 1980) showed that the recency effect in free recall experiments can persist even when the delay

[13] One might have chosen to implement a scanning model in OSCAR by assuming that the oscillators can run backwards from the time of test—stimuli would be activated sequentially as the state of the oscillators matches the state during study of each list item. While this approach would no longer be ahistorical, we would still expect scanning time to go up linearly with recency rather than logarithmically as in the present approach.

[14] In some sense, the distinction between short-term and long-term memory can be understood as a strong assertion that memory representation has a characteristic scale defined by the capacity of short-term memory.

between the last item and the test is long enough to eliminate the recency effect in delayed free recall if the spacing between the list items is increased. In addition to recency, free recall also manifests a contiguity effect. When a word is recalled, the next word recalled tends to come from a nearby list position (Kahana, 1996; Kahana, Howard, & Polyn, 2008). In analogy to the recency effect, there is also a long-term contiguity effect that is observed even with long delay intervals between the items (Howard & Kahana, 1999; Howard et al., 2008; Kiliç et al., 2013; Unsworth, 2008).

Here we focus on describing the recency effect, as measured by the probability of first recall (PFR) and the contiguity effect, as measured by the conditional response probability as a function of lag (lag-CRP) with special attention to capturing their qualitative behavior across time scales. To empirically calculate the PFR, one simply takes a serial position curve for the first word that the subject recalls in each trial. To calculate the lag-CRP, one counts the number of times that recall transitions of a certain lag are observed and divides that number by the number of times correct recall transitions of that lag could have been observed given the serial position of the just-recalled word and the previous recalls (Howard, Addis, Jing, & Kahana, 2007; Howard & Kahana, 1999; Kahana, Howard, Zaromb, & Wingfield, 2002). The "lag-CRP" curve shows remarkable consistency across experiments and subjects (Healey & Kahana, 2013; Kahana et al., 2008) and is not simply a confound of serial position effects (see especially Farrell & Lewandowsky, 2008; Howard, Sederberg, & Kahana, 2009) nor correlated encoding processes (Howard, Venkatadass, Norman, & Kahana, 2007; Kiliç et al., 2013; Siegel & Kahana, 2014).

We describe PFR and lag-CRP curves using an account very similar to that offered by TCM, but with the scale-invariant representation of temporal history playing the role of the temporal context vector. In TCM, when an item is studied, an association is formed between the current state of context at that moment. In TCM, at each recall attempt, the current state of context is used as a probe. Each potential recall is activated to the extent its encoding context matches the probe context. In TCM, when an item is remembered, it causes a partial reset of the state of the temporal context vector to the state in which that item was encoded—a "jump back in time." Here, we assume that each potential recall is activated to the extent the state of history when it was encoded matches the history at the time of test, which functions as a probe. We assume further that remembering an item causes partial recovery of the state of $\mathbf{T}$ available when the item was encoded. To describe the PFR, we use the current state of the temporal history at the time of the free recall test as a cue (Shankar & Howard, 2012). In contrast, in calculating the lag-CRP we assume that the delay and/or the process of averaging over many retrieval attempts renders the state of $\mathbf{T}$ prior to retrieval of the recalled item ineffective in generating the contiguity effect (see Farrell & Lewandowsky, 2008; Howard, Sederberg, & Kahana, 2009). The state of history used as a cue in modeling the contiguity effect is a mixture of input caused by the retrieved item itself and a "jump back in time" that recovers the previous state of temporal history. Appendix B describes the calculations of these quantities in detail.

Given the activation of each of the list items based on their match, we still need a way to map these activations onto an observable probability of recall. Let p_i be the match between state of history of the ith word in the list and state of history used as a probe. For simplicity we assume that the probability of recalling item i in a list of ℓ items is given by

a simple rule:

$$P_R(i) = \frac{p_i + c^b}{\sum_{j=1}^{\ell} p_j + c^b}, \tag{2}$$

where b is a free parameter specific to the free recall applications. $b = 2$ throughout this paper. The parameter c provides a way to take into account associative cues such as semantic similarity that do not depend on the lag between the just-recalled item and potential recalls. In calculating the PFR, we set $c = 0$.

Scale-invariant recency in the PFR. Because list items are encoded in their state of $\mathbf{T}$, and because $\mathbf{T}$ changes gradually over time, the match between the state of the history and the encoding history of each list item demonstrates a recency effect. Because $\mathbf{T}$ changes in a scale-invariant way and because the retrieval rule Eq. 2 describes a power law, this simple model predicts recency effects that are precisely scale-invariant. In the continuous distractor procedure, the subject experiences some type of stimulus, either a word or a distractor, at each moment in order to prevent rehearsal. Although the calculation is somewhat involved (see Appendix B), the result is quite simple to appreciate. Under these circumstances, if a word is presented every $\delta\tau$ seconds, it turns out that a word presented n time steps in the past will have a match of

$$p_n \simeq n^{-1}(\delta\tau)^{-2} \tag{3}$$

between its encoding history and the history at the time of test.[15] All that remains in order to apply Eq. 2 to free recall experiments is to plug in the appropriate delays between the list items and the time of test (see Eq. B6). This simple framework is sufficient to account for the basic properties of the recency effect in free recall measured in the PFR.

Property 1: A delay at the end of the list decreases the sharpness of the recency effect The top left panel of Figure 6 shows the PFRs from a free recall experiment (Polyn, et al., unpublished data) where subjects recall lists of 16 items with different end of list delays of 0, 8, and 16 seconds. In the top right panel, Eq. B6 is plotted with delays set appropriately for the three conditions.

Property 2: Increasing the gap between the words increases the sharpness of the recency effect. The bottom left panel of Figure 6 shows the PFRs of a continuous distractor free recall experiment (Howard & Kahana, 1999) where subjects recalled lists of 12 words. The conditions had a fixed retention interval of 16 s and inter-item intervals of 2, 4, 8, or 16 s. In the bottom right panel, Eq. B6 is plotted with delays set according to the experimental values. Although the numerical similarity to the data is not particularly strong, the qualitative pattern is correct.

Property 3: The recency effect persists across multiple time scales. The scale-invariance of $\mathbf{T}$ enables the model to generate recency effects across a broad variety of time scales. If nothing in the performance function (e.g., Eq. 2) disrupts the scale-invariance,

[15] Notice that $(\delta\tau)^{-2}$ will factor out when plugged into Eq. 2 with $c = 0$. Because the PFR is independent of the spacing between the items $\delta\tau$, the behavioral model is precisely scale-invariant. When $c \neq 0$, $\delta\tau$ does not factor out and the model is not scale-invariant.

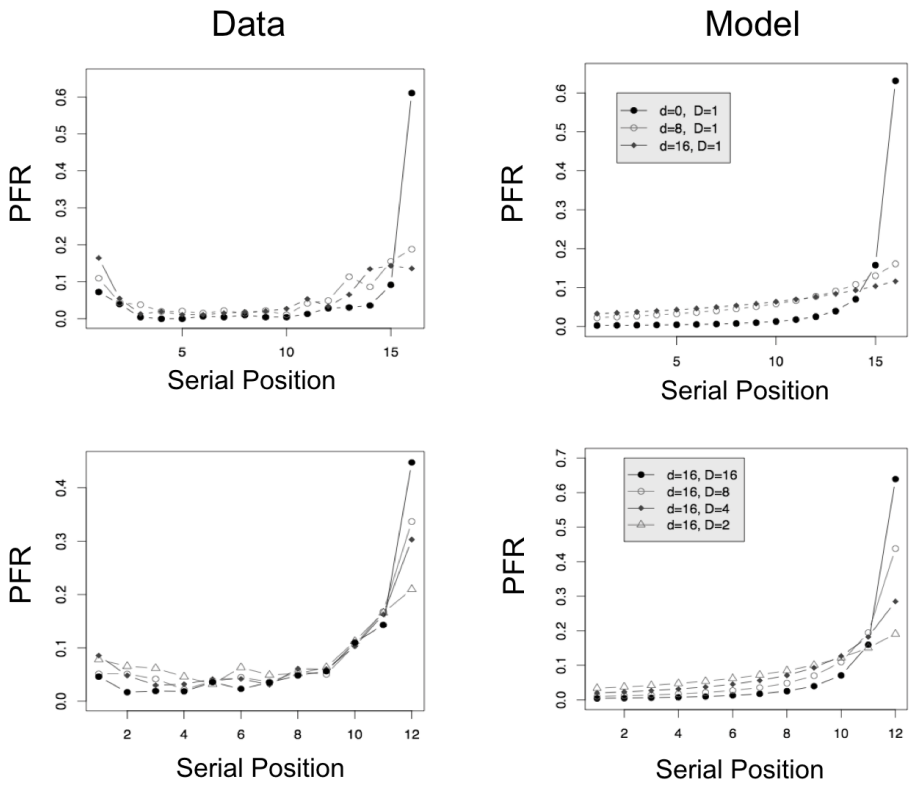

Figure 6. The recency effect in free recall depends on the relative spacing of the list. Probability of first recall (PFR) from two experiments (Polyn et al., unpublished, top; Howard and Kahana, 1999, bottom). In the Polyn et al., data (top), the duration of the filled distractor interval at the end of the list, d, is manipulated. The recency effect decreases as the delay interval increases. In the Howard and Kahana data (bottom), the duration of the distractor interval between list items, D, is manipulated while the delay prior to test is kept constant at 16 s. Increasing the delay between items enhances the recency effect. On the right side, the corresponding experiments are modeled using Eq. B6 with $b = 2$. Note the change of scale on the bottom row.

models built using $\mathbf{T}$ predict precisely scale-invariant recency. The recency effect over time-scales of seconds to dozens of seconds is already demonstrated in Figure 6. Final free recall experiments have demonstrated recency effects over even longer time scales. In final free recall experiments, subjects study and recall a number of lists. At the end of the session, subjects are asked to free recall all the words from all the lists in the order they come to mind. In final free recall experiments, the recency effect is observed spanning across several lists over several hundred seconds (Glenberg et al., 1980; Howard et al., 2008; Tzeng, 1973).

Figure 7 shows the recency effect, as measured by the PFR, from immediate and final free recall in Howard et al. (2008). In this experiment, lists of 10 words were followed by an immediate test. The PFR shows a strong recency effect extending over the last several items (top). The PFR across lists is shown in the bottom row of Figure 7. Again, there is a strong recency effect across the last several *lists* of items. We described these findings using Eq. B6 for a representative item from each list (see Appendix B for details). As for

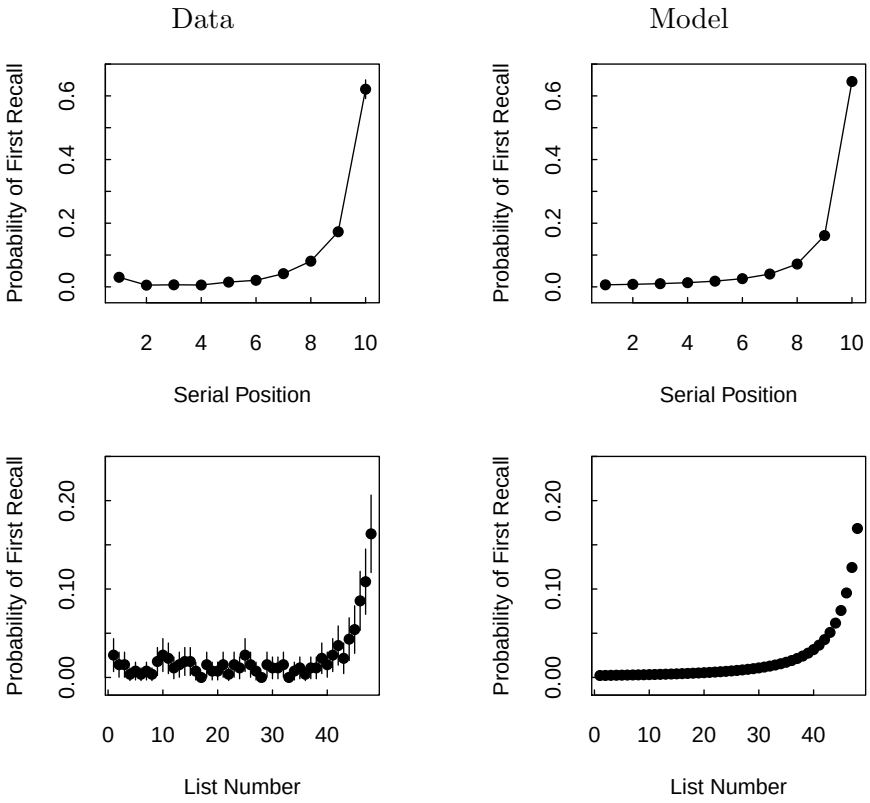

Figure 7. Evidence suggesting scale-invariant recency in free recall. Subjects performed immediate free recall on lists of ten items. After 48 such lists, they recalled all the items from all the list. Within-list (top) and across-list (bottom) probability of first recall (PFR) functions. Data from Howard, Youker & Venkatadass (2008) is shown on the left. Eq. B6 with $b = 2$ and delay parameters set to experimental values is shown on the right.

all the model predictions for free recall data, we fixed $b = 2$. For immediate free recall, we took $d = 1$ and $D = 1$; for final free recall we took $d = 250$ and $D = 49$, consistent with the values of the experiment.

Contiguity. We consider only the contribution of the just-recalled word to $\mathbf{T}$ in calculating the probability of recalling each of the candidate words from the list. This is certainly a simplification. In addition to the most recently-recalled word, the recency effect can also contribute to subsequent recalls affecting the shape of the lag-CRP (see Howard, Sederberg, & Kahana, 2009, for a thorough discussion). Moreover, there is also evidence that several previous recalls, not just the most recent one, can affect the lag-CRP (Lohnas & Kahana, 2014). However, the lag-CRP with its characteristic shape is observed in free recall even when the recency effect is absent. Moreover, the contribution of remote recalls appears to have the same qualitative properties as the contribution of the immediately-previous recall. Ignoring recency and compound cuing enables us to focus on the essential properties of the contiguity effect.

One of the notable features of the lag-CRP in free recall is the near-ubiquitous asymmetry. Across a variety of studies the lag-CRP corresponding to adjacent forward transitions is higher than the lag-CRP corresponding to adjacent backward transitions (Kahana et al., 2008). We account for this asymmetry in a way analogous to the account TCM provides. Briefly, when a word is recalled, it has two effects on the temporal history. First, because this word is the same as a word from the list, as this word enters the history and gradually moves back in time, the history resembles the encoding history of words in the list that followed the inital presentation of the word. Consider the list A B C D E. The encoding history of D contains the word C one time step in the past; the encoding history of E contains C two time steps in the past. In the time after C is recalled, the history again includes C. This match provides a basis for forward associations between C and words that followed its original presentation. In addition, when a word is recalled it can also retrieve the prior state of history. The history when C was originally presented matches the encoding history of D more than it matches the encoding history of E. In addition, the jump back in time results in a cue that also supports a backward association; the history recovered by C matches the encoding history of B moreso than it matches the encoding history of A. Appendix B works out the details of the calculation of the match. A free parameter γ controls the relative contribution of these two components, with $\gamma = 0$ corresponding to no recovery of the previous state of history.

There are two additional considerations prior to examining model predictions. First, the effectiveness of repeating an item as a cue for subsequent recalls changes in the moments after τ_r when the stimulus is repeated. For simplicity, we assume that the state a non-zero time r after repetition of the word, *i.e.,* at time $\tau_r + r$, is the cue for retrieval of the next word. Second, as discussed above, there are a great many reasons why one word would cue another in a way that does not depend on their temporal contiguity. The c parameter in Eq. 2 accounts for these factors. In the illustrations of model predictions that follow, we set $\gamma = 0.9$, $r = .75$, corresponding to 750 ms if $\delta\tau = 1$ s, and $c = 8 \times 10^{-5}$. As in the PFR calculations, $b = 2$.

Figure 8 shows the lag-CRPs of the final free recall in the experiment of Unsworth (2008).[16] Unsworth (2008) presented subjects with 10 lists of 10 words each. Each list was presented at a rate of 1 word per second followed by a 16 second distracting math task, subsequently followed by a 45 second recall period for that list. The next list was presented immediately following the end of the previous recall period. We take the time interval between the study of two successive lists as 71 seconds. During the final recall, whenever a successively recalled pair comes from the same study list, it is grouped as a within list transition. The CRP of the within-list transitions is plotted in the top left panel. Whenever the successively recalled pairs were from different lists, they are grouped as across-list transitions, meaning an across-list transition of zero is not possible in this analysis. The lag-CRP of the across-list transitions is plotted in the bottom left panel. Note that the empirical data are not strictly scale-invariant. If they were, the within- and across-list curves would perfectly align. In addition to the change in the steepness of the curves, there is also a change in the asymmetry apparent in the data.

[16] Howard et al. (2008) conducted similar contiguity analyses of FFR data and found qualitatively similar results. However, in that experiment there was residual recency in the contiguity effect observed in FFR, so that the simplifiying assumptions used in Appendix B do not apply.

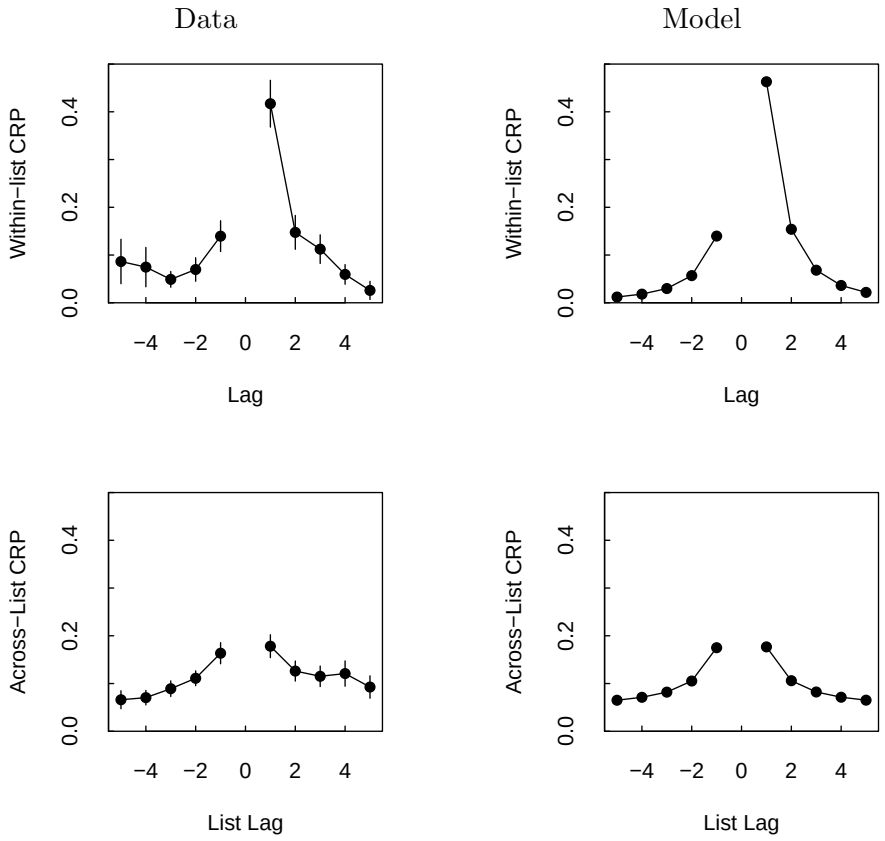

Figure 8. The contiguity effect persists across time scales. After a word is recalled from serial position i, the lag-CRP estimates the probability that the next recalled word comes from position $i + \text{lag}$. Within- and across-list lag-CRP were measured during final free recall by focusing on transitions between words from the same list and between words from different lists respectively. The across-list lag-CRP (bottom) estimates the probability of a transition from *list* i to list i+list lag. Left: Data from Unsworth (2008). Right: Eq. 2. See text for details.

Equation 2 is plotted on the right side of Figure 8. The model captures the qualitative aspects of the data. Both within- and across-list contiguity effects exist in both the forward and backward directions. Within-list transitions are asymmetric favoring forward transitions. Across-list transitions are more nearly symmetric and lower in amplitude. The behavioral model deviates from scale-invariance, despite utilizing a scale-invariant representation of internal time, because the behavioral model fixes a scale. First, the parameter c fixes a scale, resulting in the behavioral model's prediction of a more shallow across-list contiguity effect. The model captures the change in asymmetry across time scales because r, the time after recall of the preceding word, also fixes a scale. For the within list transitions, r is similar to the time interval between list items (one second) and we see a strong forward asymmetry. When r is small compared to the study time between items, the lag-CRP will be more nearly symmetric because the temporal history when the stimulus was initially presented will dominate.

Discussion. Modeling of recency in the PFR and contiguity as measured by the lag-CRP is a requirement for a free recall model, but it does not constitute a complete model of the free recall task. Using the temporal history $\mathbf{T}$ as a generalization of temporal context does not sacrifice the ability to describe recency or contiguity effects, showing performance at least comparable to the first presentation of TCM (Howard & Kahana, 2002). In the following years, several treatments of the details of free recall dynamics have been presented (e.g., Howard, Sederberg, & Kahana, 2009; Lohnas & Kahana, 2014; Polyn et al., 2009; Sederberg et al., 2008; Siegel & Kahana, 2014). These models elaborate the basic associative engine described in the Howard and Kahana (2002) paper by considering the effect of compound cueing, resampling, and a detailed moment-by-moment retrieval process. These modeling efforts have shown that the basic associative engine of the original paper can be used to account for the persistence of recency across multiple retrieval attempts, latency differences in recalls as a function of output position, lag effects, and the effects of semantic similarity on retrieval dynamics, among other properties. The only substantive difference between the treatment of recency and contiguity presented here using a scale-invariant history and the treatment of recency and contiguity in the Howard and Kahana (2002) paper using the temporal context vector is the fact that the history is genuinely scale-invariant whereas the temporal context vector is not (Howard, 2004). This gives the present treatment the ability to account for recency and contiguity effects over much longer scales without changing parameters.

Temporal mapping

In temporal mapping experiments (e.g., Cole et al., 1995; Barnet, Cole, & Miller, 1997; Savastano & Miller, 1998; Arcediano & Miller, 2002; Arcediano, Escobar, & Miller, 2003), subjects both learn temporal relationships between pairs of stimuli and infer temporal relationships between stimuli that were not actually experienced close together in time. Miller and colleagues have developed the temporal encoding hypothesis (Arcediano & Miller, 2002; Matzel et al., 1988; Savastano & Miller, 1998) to describe learning. The temporal encoding hypothesis has two components that are particularly relevant for the model of internal time developed here. First, Miller and colleagues argue that the temporal relations between stimuli forms an essential and unavoidable part of the learning event. Second, they argue that learners can integrate disparate learning events into a coherent temporal map by aligning different time-lines on a common stimulus. We model core results from the temporal mapping paradigm using matching and the ability to jump back in time to a previous state of history. Matching of a state of temporal history enables an appropriately timed response that reflects the temporal relationship between the CS and US; jumping back in time provides a mechanism for the integration of temporal relationships across learning episodes.

To make temporal mapping more concrete, let us describe a specific experiment. Cole et al. (1995) trained rats to associate a 5 s CS1 with a US (shock). In one condition, the time between offset of the CS1 and the onset of the US was 0 s (Figure 9a, top). In the other condition, the time between the offset of CS1 and the US was 5 s (Figure 9a, bottom). Let us refer to these as the 0 s and 5 s conditions, respectively. After training the CS1-US association, a second-order association was formed between CS1 and another 5 s CS2. In both conditions, the onset of CS2 immediately followed the offset of CS1 (Figure 9). In

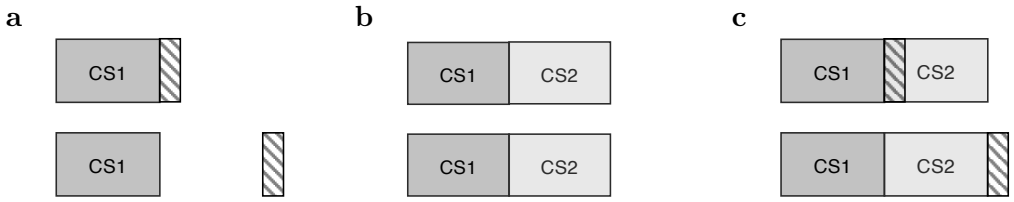

Figure 9. Procedure of a typical temporal mapping experiment (after Cole, Barnet & Miller, 1995). Both conditions involve two conditioned stimuli. The conditions differ in the duration of the gap between CS1 and the US during an initial phase of learning. After learning, one can evaluate both trace conditioning (the conditioned response to CS1) and second order conditioning (the conditioned response to CS2). **a.** In the first phase of training, subjects in the 0 s condition (top) received CS1 with the shock US presented immediately after CS1 offset. Subjects in the 5 s condition (bottom) received the US five seconds after offset of CS1. There was more robust conditioned responding (CR) to CS1 in the 0 s condition. **b.** In a second phase of training, both groups received second-order conditioning between CS1 and CS2. **c.** The presumed integrated representation according to the temporal mapping hypothesis after both phases of training. CS2 evoked a larger CR in the 5 s condition than in the 0 s condition.

neither condition did CS2 ever cooccur with the US. The first finding was, not surprisingly, that the CR to the CS1 was stronger in the 0 s condition than in the 5 s condition. If the relationships learned between the stimuli were atomic associations, we would expect the second order conditioning to CS2 would also be stronger in the 0 s condition than in the 5 s condition. After all, the association from CS2 to the US must be mediated by the association from CS1 to the US. However, exactly the opposite was observed: the CR to CS2 was greater in the condition where the CR to CS1 was smaller. This result makes no sense from the perspective of simple associative strength. The temporal coding hypothesis (Matzel et al., 1988; Savastano & Miller, 1998) of Miller and colleagues reconciles these findings as follows. Note that if the two learning episodes were integrated into a single coherent representation aligned on the CS1 (as in Figure 9c), then the CS2 would not predict the onset of the US in the 0 s condition. In the 5 s condition, CS2 strongly predicts the onset of the US when the two learning episodes are aligned on presentation of CS1.

A behavioral model must have two basic properties in order to account for the temporal mapping phenomenon. One is that the temporal relationship between stimuli, rather than a simple scalar associative strength, is learned. Second, there must be some mechanism for integrating different experiences into a coherent synthetic representation. The representation of internal past and future time satisfies the first constraint. The ability of a stimulus to recover previous states of temporal history satisfies the second constraint (see also Howard et al., 2005; Howard, Jing, et al., 2009; Rao & Howard, 2008). When CS2 is given as a probe if it recovers the state of history from the first phase of learning, the state of $\mathbf{T}$ it recovers doesn't just provide a general sense that it has something to do with CS1. Rather, it recovers information specifying that CS1 was presented 5 s in the past. This recovered history provides an effective cue for the US to the extent it matches the state of history when the US was originally experienced. As time passes after the presentation of CS2, the representation of CS1 drifts further into the past, changing the match with the

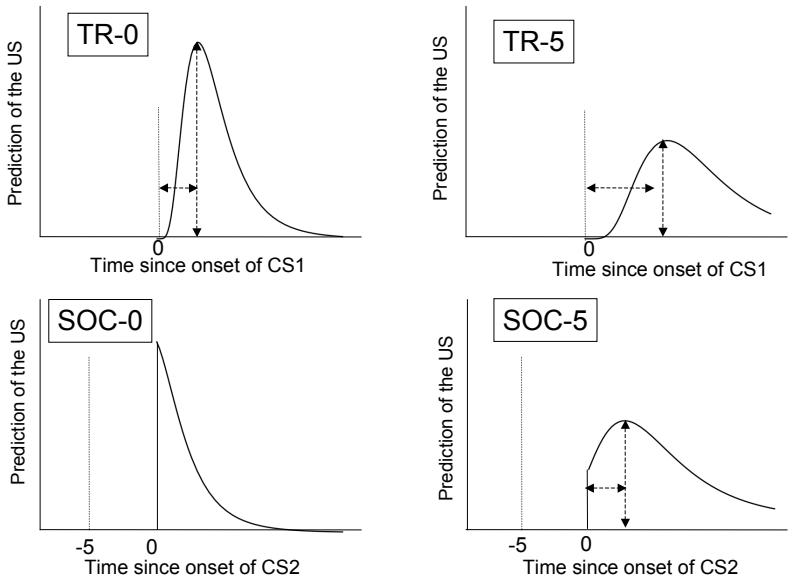

Figure 10. Schematic illustration of the match to the encoding history of the US in the time after the onset of CS1 (trace conditioning, TR) or CS2 (second order conditioning, SOC) in the two conditions described in Figure 9 (0 s delay or 5 s delay between CS1 and the US). The top row shows the match following CS1 (trace, TR). The bottom row shows the match following CS2 (second order conditioning, SOC). The left column shows results for the 0 s condition (top row of Figure 9); the right column shows results for the 5 s condition (bottom row of Figure 9). In the second order conditions, even though CS2 was not explicitly paired with the US, the match with the US is nonzero because CS2 recovers a temporal history that includes CS1 5 s in the past, the delay between the onsets of CS1 and CS2 during training (see Figure 9b). The temporal profile of the match following CS2 is the same as the match following CS1 with a time shift of 5 seconds.

US (Figure 10)

In the Cole et al. (1995) experiment the CR is measured as the time it takes for the rat to drink a threshold quantity of water after the CS is presented. To mirror this measure in a simple behavioral model, we assume that the rate of water drinking is exponentially suppressed as a function of the normalized match between the history after presentation of the CS and the encoding history of the US.[17] Since the rat has been trained for a long time, we assume that the match is normalized to a peak value of 1 in trace conditioning. For the second order conditioning, the match is reduced by a factor of γ. Figure 11 shows

[17]That is, at each moment τ after the CS, the rate at which the rat drinks is given by $e^{-dp(\tau)}$, where d is a free parameter and $p(\tau)$ is the match between the encoding state of the US and the current state of $\mathbf{T}$ at time τ. We then solve

$$\text{Threshold} = \int_0^L e^{-dp(\tau)} d\tau$$

to find the value of L when the threshold amount of water is consumed.

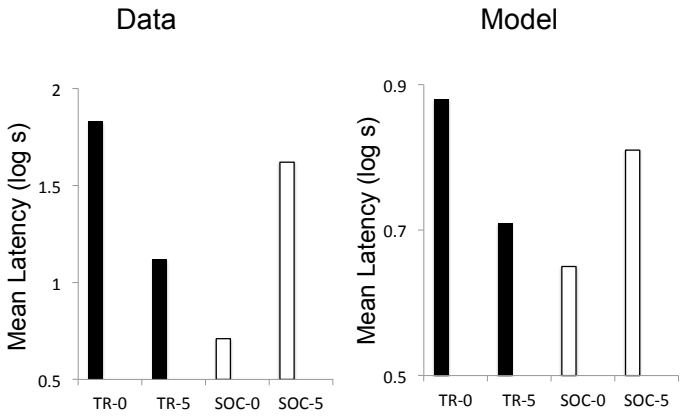

Figure 11. Results from a temporal mapping experiment. Left: Data from Cole, Barnet & Miller (1995). The CR after training to CS1 (the TR conditions, black bars) or CS2 (the SOC conditions, light bars) separated by the condition (0 or 5), corresponding to the delay in seconds between the CS1 and the US in the first phase of learning (see Figure 11a). The CR in each condition is the log of the the mean time taken by the rats to drink water for five cumulative seconds. Fear of the shock US disrupts drinking. Right: Predictions of a simple behavioral model in which the rate of drinking is suppressed by the degree to which the state of history matches the encoding history of the US (see text for details).

the predictions of this simple behavioral model with the threshold set to 3, the d parameter set to 5/3, and γ set to 0.5.

Discussion. Temporal mapping is not limited to the findings from the Cole et al. (1995) experiment. For instance, if the order of the training phases is reversed, training the relationship between CS1 and CS2 prior to learning the relationship between CS1 and the US, the same qualitative results are obtained (Experiment 2, Cole et al., 1995). The framework described here would predict the same outcome even if the stages of training were reversed. Other variants of the paradigm (Arcediano et al., 2003; Arcediano, Escobar, & Miller, 2005; Barnet et al., 1997) may be considerably more challenging, requiring a significant elaboration of the model. For instance, in Experiment 2 of Arcediano et al. (2003), rats first learned that CS2 preceded CS1 by five seconds. In a second phase of learning, they learned that the US preceded CS1 by four seconds. Integrating the phases of learning by aligning on CS1, the only element common to both phases of learning, would result in the offset of CS2 immediately preceding the US onset. Indeed, there was robust conditioned responding to CS2 relative to control conditions. Because the US is never preceded by either of the conditioned stimuli, this experiment constitutes a challenge to mechanistic models that require a match to the encoding history of the US.

Neurophysiological evidence for a representation of internal time

If a scale-invariant time-line of recent experience is at the center of a wide variety of cognitive functions, then we would expect to see signatures of this representation in a wide variety of brain regions. There are three primary properties predicted for this type of

representation. First, a scale-invariant representation of history predicts long-range correlations caused by the to-be-remembered stimuli. Second, it should be possible to recover states of this gradually-changing representation. The jump-back-in-time was essential for accounting for the contiguity effect in episodic memory as well as the integration of learning episodes in the temporal mapping experiments. Third, the framework requires not only that the representation change over time in a way that is sensitive to the stimuli, but also that the representation contain explicit information about when stimuli were experienced. We review evidence for all three of these predictions (see also Howard & Eichenbaum, 2013).

Gradually-changing stimulus-specific representation. A scale-invariant time-line of recent experience would cause neurophysiological measures to be temporally autocorrelated over long time scales, in a way that both depends on the stimuli presented and also correlates with performance in behavioral tasks. We discuss evidence for these properties from the prefrontal cortex (PFC) and medial temporal lobe (MTL) with special attention to the JOR task.

Recent evidence suggests that JOR performance is supported by distributed representations in the PFC and MTL that change gradually over time. Jenkins and Ranganath (2010) observed BOLD response using fMRI while subjects encoded a set of stimuli for a subsequent absolute JOR task. The multivoxel pattern of activity in the PFC during study changed gradually across the list. The rate of change of the pattern in one region, the rostrolateral PFC, predicted subsequent performance on the JOR task. Ezzyat and Davachi (2014) examined human JOR performance for stimuli (pictures of objects) that were paired with a context picture of a scene while recording multivoxel patterns of activity in the MTL. The distance between multivoxel patterns in the hippocampus predicted subjects' judgements for pairs of probes studied in different contexts, as if the distance in the subjects' neural response was driving the rating of distance in the behavioral response. Manns, Howard, and Eichenbaum (2007) measured the firing rate of ensembles of neurons recorded simultaneously from the hippocampus of rats performing a JOR task in which they studied lists of odors. They found that the distance between the population vectors corresponding to study of the list odors increased with temporal distance both within and across lists. The ensemble of neurons in the hippocampus continued changing gradually over more than a thousand seconds (see also Hyman, Ma, Balaguer-Ballester, Durstewitz, & Seamans, 2012; Kim, Ghim, Lee, & Jung, 2013; Mankin et al., 2012).

The foregoing results do not provide strong evidence that the gradually-changing representation reflects memory for the to-be-remembered stimuli; those results could simply be due to stochastic noise rather than memory for the identity of the presented stimuli. There is also evidence, from tasks other than the JOR task, that indicates that the MTL and PFC contain a gradually changing representation that retains information about past stimuli. Schoenbaum and Eichenbaum (1995a, 1995b) presented rats with a series of odors. They observed that some neurons in the rodent PFC responded on a trial to the identity of the stimulus on the preceding trial. In the MTL, Yakovlev, Fusi, Berman, and Zohary (1998) showed that neurons in the monkey perirhinal cortex, a cortical region within the MTL, showed stimulus-specific activity during a delayed-match-to-sample task that persisted across trials. Hsieh, Gruber, Jenkins, and Ranganath (2014) presented human subjects with series of pictures of objects. They found that the multivoxel pattern of ac-

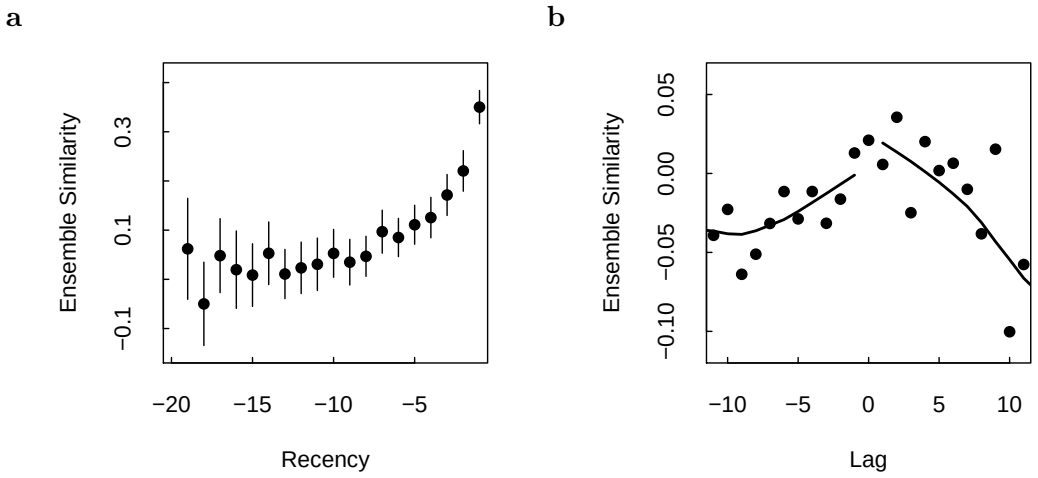

Figure 12. Neural recency and contiguity effects. Multiple neurons were recorded from human MTL during performance of a continuous recognition task. **a.** The vector of firing rate across cells averaged over the 3 s each stimulus was presented was compared to the vector from previous presentations. Comparisons were restricted to stimuli within the same block of stimuli and excluded comparisons to the same stimulus. The ensemble changed gradually over macroscopic periods of time, here up to about 20 s. **b.** The population vector when a stimulus was repeated was compared to the vector observed during presentation of the neighbors of its original presentation. After statistically eliminating the contribution due to recency, these analyses revealed a contiguity effect. After Howard, Viskontas, Shankar & Fried (2012).

tivity in the hippocampus changed gradually across the presentation of multiple stimuli in a way that reflected the identity of the preceding stimuli. Other recent fMRI studies suggest multivoxel patterns of activation in the MTL form a temporal context that drives performance in episodic memory tasks (Gershman, Schapiro, Hupbach, & Norman, 2013; Chan et al., in preparation).

Neural evidence for jumping back in time during performance of episodic memory tasks. In several of the cognitive applications described here, we assumed that prior states of the time-line can be recovered. In free recall, this accounted for the contiguity effect. Recovery of previous states of internal time was also essential for bridging across disparate experiences in the temporal mapping application. Previous work has argued that one of the functions of the MTL, and the hippocampus in particular, is to enable this jump-back-in-time (Howard et al., 2005). Although there is not yet definitive evidence that the brain is able to recover prior states of a temporal history, we discuss several recent studies in humans that are suggestive of this hypothesis.

Zeithamova, Dominick, and Preston (2012) showed evidence suggesting a jump-back-in-time using fMRI data. They had subjects learn simultaneously-presented pairs of pictures. The pairs shared an overlapping stimulus, i.e., A-B and B-C. Although A and C were not learned at the same time, they were both experienced in the context of B. Previous work has shown that bridging associations between A and C depend on the integrity of

the hippocampus (Bunsey & Eichenbaum, 1996; Greene, Gross, Elsinger, & Rao, 2006). In the Zeithamova et al. (2012) study, the stimuli were chosen from different visual categories such that the category could be distinguished using multivoxel pattern classification. During study of B-C, the cortical pattern associated with the category of stimulus A was preferentially activated even though A was not physically present. The degree to which A was activated was correlated with the degree to which the anterior MTL, including the hippocampus, was activated by learning of the pairs.

Manning, Polyn, Litt, Baltuch, and Kahana (2011) found evidence suggesting a jump-back-in-time supports the contiguity effect in free recall. They recorded intracranial EEG from patients with epilepsy during study and free recall of a list of words. They found that the pattern of oscillatory components changed gradually over time during study of the list. Notably, the pattern of activity just before recall of a word from the list resembled the pattern during study of neighbors of the original presentation of the about-to-be-recalled word. This similarity showed a contiguity effect and was greater for subjects who showed a larger behaviorally-observed contiguity effect.

Howard, Viskontas, Shankar, and Fried (2012) reanalyzed single-unit recordings from the MTL of epileptic patients performing a continuous recognition task on images (Viskontas, Knowlton, Steinmetz, & Fried, 2006). The population vector across neurons after the presentation of each stimulus was computed. Population vectors changed gradually across at least 20 s (Fig. 12a), demonstrating a neural recency effect. When a stimulus was repeated as an old item, the gradually-changing part of the ensemble resembled the neighbors of the original presentation of that stimulus (Fig. 12), exhibiting a neural contiguity effect.

While each of the studies described here have some limitations, their limitations are complementary, so that taken as a group they present a much stronger story than any study in isolation. Although the Manning et al. (2011) result could be due to correlated retrieval strategies, neither the Howard et al. (2012) result nor the Zeithamova et al. (2012) result are susceptible to that concern. While the Zeithamova et al. (2012) experiment involved stimuli presented simultaneously, so that the finding does not require recovery of a gradually-changing memory state, both of the other two papers were able to establish a gradually-changing representation. While the Howard et al. (2012) paper relied on statistically isolating the contiguity effect from the recency effect, the methods of both of the other two papers were such that there was little or no recency effect and the neural contiguity effect could be directly measured.

Evidence for a representation of an internal time-line. Any number of methods would result in a representation that changes gradually over time—for instance an ahistorical strength model (Howard & Kahana, 2002), a set of coupled oscillators (Brown et al., 2000), or stochastic noise (Estes, 1955; Mensink & Raaijmakers, 1988). A short-term memory buffer or a decaying strength vector would cause a gradually-changing representation that is specific to recent stimuli; if the buffer contents or temporal context vector were recovered by a repeated stimulus this would account for a jump back in time. However, an internal time-line should not only change gradually over time, but also contain information about what stimulus was experienced when. The key prediction of this form of conjunctive what-and-when coding is that cells should respond to their preferred stimulus not immediately,

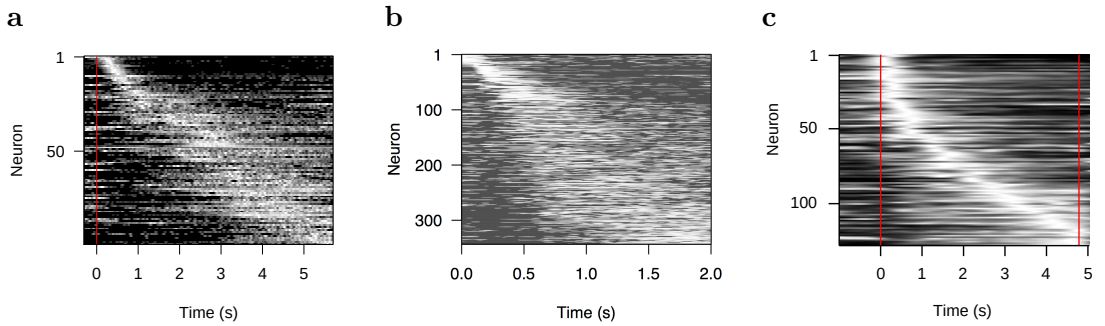

Figure 13. Temporal coding in diverse brain regions. In each panel, the average firing rates of many cells from a particular brain region during a delay interval are shown. Each row is the firing rate of one neuron averaged over trials with white indicating high firing rate and black indicating low firing rate. The neurons are numbered such that cells that fire earlier during the delay are at the top of the figure and neurons that fire later in the delay are at the bottom. Many cells fire for some circumscribed interval of time within the delay. The times at which those cells peak are distributed smoothly across the delay interval. Cells that fire later in the delay show a wider spread in their activation. **a.** Neurons in the rodent hippocampus during the delay of an object memory task. Re-analysis of MacDonald, et al., (2011). **b.** Medium spiny neurons recorded from the ventral striatum of monkeys during the delay of a classical conditioning task. Replotted from Adler, et al., (2012). **c.** Neurons from the rodent medial prefrontal cortex during the delay period of an interval discrimination task. Re-analysis of Kim, et al., (2013). After Tiganj, Kim, Jung, and Howard (in preparation).

but after some characteristic time (Fig. 2, see also Shankar & Howard, 2012; Howard et al., 2014). MacDonald, Lepage, Eden, and Eichenbaum (2011) observed that neurons in the rodent hippocampus fired at circumscribed periods of time after the beginning of a delay period (Figure 13a, see also Pastalkova, Itskov, Amarasingham, & Buzsaki, 2008). These cells have been dubbed "time cells" because they fire in circumscribed regions of time analogous to hippocampal "place cells" which fire in circumscribed regions of allocentric space (O'Keefe & Dostrovsky, 1971; Wilson & McNaughton, 1993; see also Pastalkova et al., 2008). Some hippocampal time cells have been shown to fire differentially in response to the stimulus that precedes a delay (Gill, Mizumori, & Smith, 2011; MacDonald et al., 2011; MacDonald, Carrow, Place, & Eichenbaum, 2013; Pastalkova et al., 2008), thus providing a record of what happened when. Howard et al. (2014) confirmed several qualitative predictions of **T**, including the spread in time fields that code for longer durations (see also Kraus, Robinson, White, Eichenbaum, & Hasselmo, 2013) and also the asymmetry in their time fields.

The foregoing suggests that cells in the hippocampus, a region critically important in episodic memory, show qualitative properties consistent with those we would expect of **T**. If the representation of an internal time-line is accessed for a variety of tasks in memory and timing, then, to the extent those different tasks depend on different brain regions, we would also expect to see the signature of the distributed representation of temporal history in many different brain regions. Although this literature is not nearly as developed as the hippocampal literature, some recent findings suggest an internal time-line in the striatum and the medial PFC.

Adler et al. (2012) showed evidence for stimulus-specific time cells in the monkey striatum during performance of a Pavlovian conditioning task. They recorded from medium spiny neurons in the putamen, a subregion of the dorsal striatum, while monkeys were presented with simple stimulus-response pairings (Adler, Finkes, Katabi, Prut, & Bergman, 2013, later reported similar results for medium spiny neurons in other regions within the striatum). Different stimuli predicted different outcomes. Figure 13b shows the response of each of the medium spiny neurons as a function of time during the delay between the stimulus and the outcome. The medium spiny neurons show several properties predicted for cells participating in $\mathbf{T}$. First, different cells respond at different latencies after the stimulus. Second, the temporal spread in the firing of cells increases with the latency at which they respond. Finally, although Adler et al. (2012) did not analyze the individual stimuli separately, they did show that medium spiny neurons responded differently on trials with different outcomes.

Kim et al. (2013) showed evidence for time cells in the rodent PFC during performance of a temporal discrimination task. The rats were required to wait in a circumscribed region of a T-maze for some delay. If the delay was shorter than a criterion, a left turn resulted in reward; if the delay was longer than a criterion, a right turn resulted in reward. Figure 13c shows the firing rate of a subset of the neurons that fired during the delay interval. Again, this set of neurons showed peaks of firing that covered the entire delay interval, and firing fields that spread out with the delay. Intriguingly, the distribution of the peak firing times was not uniform (note the curved rather than linear ridge in Figure 13c), qualitatively consistent with a Weber-Fechner spacing. The width of the temporal receptive fields was broader for cells that peaked later in the interval (Z. Tiganj, personal communication 7/11/2014), qualitatively consistent with decrease in temporal accuracy with delay in $\mathbf{T}$.

While many additional experiments and analyses will be necessary to determine if the "time cells" in various regions are quantitatively consistent with the properties of cells supporting $\mathbf{T}$, it is striking that such similar results are observed in the hippocampus, the striatum, and the medial PFC. The hippocampus is believed to be part of the declarative memory system (Squire, 1992; Cohen & Eichenbaum, 1993), which is critically important for episodic memory. The PFC is believed to be important in working memory and executive control (Braver et al., 1997; Miller & Cohen, 2001). The striatum, including the putamen, is usually understood to be part of a non-declarative habit learning memory system (e.g., Knowlton, Mangels, & Squire, 1996; Squire & Zola, 1996) and believed to be critically important in a variety of timing tasks (see Buhusi & Meck, 2005, for a review). Perhaps the differences between the different memory systems can be understood in terms of the different operations they make on a common representation of internal time.

General Discussion

We constructed a series of behavioral models using a scale-invariant representation of what happened when (Shankar & Howard, 2012, 2013). The results suggest that this representation, when accessed and exploited with appropriate operations, is sufficiently rich to describe performance in tasks from a variety of fields, including working memory, episodic memory and conditioning. We reviewed a growing body of neurophysiological evidence that shows properties similar to those predicted for the representation of an internal time-line in diverse brain regions believed to support distinct memory systems. The behavioral models

used here relied on very distinct operations. For instance, whereas the JOR application relied exclusively on scanning, the free recall application did not, instead relying on matching and the recovery of a previous state of temporal history. These findings suggest that a scale-invariant internal time-line could provide the basis for a physical theory that ties together many forms of memory.

Ordered representations of variables other than time

The method for constructing a scale-invariant internal time-line (Shankar & Howard, 2012, 2013) can be generalized to construct scale-invariant representations of variables other than time (Howard et al., 2014). In most of the behavioral experiments described here, the data do not require an intrinsically temporal representation. For instance, in JOR experiments in which stimuli are repeated at a constant rate, the time since a stimulus was presented is precisely confounded with the number of intervening stimuli. Indeed, interference accounts of forgetting have a long history. An interference-based account of forgetting can be constructed using an ordinal representation.

To construct an ordered ordinal representation of past experience, it is necessary to alter the rate at which the leaky integrators in the intermediate representation $\mathbf{t}(s)$ change.[18] If the input is a series of discrete stimulus presentations separated by various delays, then in order to construct an ordinal representation, the intermediate representation must not change except when a stimulus is being presented.[19] This causes the representation to remain constant between stimulus presentations, advancing only when a stimulus becomes available. The $\mathbf{T}$ constructed from this intermediate representation is an ordinal representation of what happened in what order, ignoring the delays between stimulus presentations. This representation still retains the property of scale-invariance. Because it is still an ordered representation, it can still support scanning.[20]

A quick survey of findings on time in memory suggests that subjects can construct and utilize both temporal and ordinal representations in different circumstances. Some results suggest that memory should be sensitive to time *per se*. For instance Hintzman (2004) presented subjects with a continuous absolute JOR task in which the items were presented for different times. The ratings of recency depended on the intervening time rather than the number of intervening items. In contrast, there is evidence that temporal gaps enhance item recognition (Morin, Brown, & Lewandowsky, 2010). In recall tasks, while there is evidence that free recall performance is sensitive to temporal delays between list items (Brown, Morin, & Lewandowsky, 2006), the recency effect is unaffected by unfilled delay intervals (Baddeley & Hitch, 1977; Murdock, 1963). In serial order memory, Lewandowsky and Oberauer and colleagues have argued that temporal gaps have little or no effect on short-term serial order memory (Lewandowsky & Oberauer, 2009; Lewandowsky, Oberauer, & Brown, 2009; Oberauer & Lewandowsky, 2008, 2013), but this conclusion is not universally supported (Lewandowsky, Nimmo, & Brown, 2008; Morin et al., 2010).

[18] More formally, it is necessary to change all the values of s in register by setting $s' = \alpha(\tau)s$.

[19] More formally, set $\alpha(\tau) = |\mathbf{f}(\tau)|$.

[20] By allowing the values of s to change in response to other variables, this framework can also be used to construct scale-invariant representations of other variables, including spatial position and numerosity (Howard et al., 2014).

Relationship to previous clock models of timing and memory

T provides a distributed representation of both what and when information. In contrast, clock models keep track of time relative to some starting point (Brown et al., 2000; Gibbon, 1977; Matell & Meck, 2004). Clock models can be divided into models that keep of track of time with a single counter *vs* models with a distributed clock signal. The present method for representing internal time is more flexible than either of these classes of models and is thus able to avoid conceptual problems that limit the applicability of clock models of both types.

In counter models a counter keeps track of a set of pulses accumulated over time (e.g., Gibbon, 1977). An advantage of counter models is that they can be easily related to drift diffusion processes (e.g., Luzardo, Ludvig, & Rivest, 2013; Simen, Balci, Souza, Cohen, & Holmes, 2011). A disadvantage of counter models is the difficulty in combining information about multiple intervals. For instance, scalar expectancy theory, a widely used model of timing and conditioning (Gibbon, 1977; Gallistel & Gibbon, 2000; Rakitin et al., 1998), encounters a serious conceptual problem that is caused by the way time is represented (Machado & Silva, 2007). Consider the following experiment. On each trial, animals are rewarded either for the first response after a short delay, say 10 s, or the first response after a long delay, say 120 s. When reward is omitted, animals show one peak of responding around 10 seconds and another peak around 120 s (Catania & Reynolds, 1968; Leak & Gibbon, 1995). If one simply averaged the counter values that obtain when a reward is delivered into a common memory store, one would expect a unimodal peak somewhere between 10 and 120 s. For scalar expectancy theory to account for this result, there must be two memory stores corresponding to the responses around the two time intervals. As Machado and Silva (2007) point out, this is conceptually problematic: the system must have some way to know which store in which to place a given observation, which requires a way to discriminate the intervals, which are identical except for their temporal content (Gallistel, 2007).

Distributed clock models are not subject to this conceptual problem. If the clock setting is distributed across the activity of many nodes, the combination of 10 s and 120 s can be a bimodal distribution with a peak around 10 s and another peak around 120 s. However, distributed clock models have other limitations. Consider how OSCAR (Brown et al., 2000) would address appropriately timed associations between a CS and US in trace conditioning. In OSCAR, the clock runs through a series of states, described by a set of oscillators of different frequencies (Church & Broadbent, 1990; Gallistel, 1990; Miall, 1989) that run independently of the stimuli that are presented. In serial recall applications, the autonomy of the memory representation has been considered to be a major advantage over chaining models in which the internal state depends on successful recall of previous list items (Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Henson, Norris, Page, & Baddeley, 1996, but see Farrell & Lewandowsky, 2004; Hulme, Stuart, Brown, & Morin, 2003; Solway, Murdock, & Kahana, 2012). Because the internal states in OSCAR are autonomous of the stimuli, one must decide to externally "restart the clock" to proceed through the same states as before. Suppose a CS is presented, starting the oscillators, and then later a US is experienced in some state of the oscillators. One could account for a CR if repeating the CS later on causes a reset of the state of the oscillators. Then, the oscillators will run forward autonomously and reach the same state they were in when the US was presented, enabling

one to predict the arrival of the US. But what if there are two CS-US pairings that have been learned? Suppose that CS1 is presented and before US1 arrives, CS2 is presented. At this stage, there are two choices. If the oscillators run independently of CS2, then US1 can be predicted, but US2 cannot. Conversely, if the oscillators are reset by CS2, then US2 can be predicted but US1 cannot.

The scale-invariant time-line used here resembles a clock model in that following presentation of a stimulus, the system runs through a series of states autonomously. However, the states caused by distinct stimuli are themselves distinct. The state with CS1 a certain time in the past need not overlap with the state with CS2 a certain time in the past. In this respect, the model of internal past time used here has much in common with spectral resonance theory (Grossberg & Merrill, 1992, 1996). Because of this, there is no difficulty in simultaneously retaining information about the separate stimuli and thus correctly predicting both US1 and US2. That is, because $\mathbf{T}$ contains conjunctive information about what and when, it can describe both autonomous sequential dynamics following a stimulus as well as account for causal relationships between memory cues and outcomes. The autonomous dynamics provide the ability to learn temporal intervals. However, the ability to learn and express causal temporal relationships between stimuli, which are observed in episodic memory tasks (Kiliç et al., 2013; Schwartz, Howard, Jing, & Kahana, 2005), sets it apart from other recent models of episodic recall (Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Grossberg & Pearson, 2008; Farrell, 2012).

The benefit of conjunctive coding of what and when comes at a cost. The number of nodes required to maintain a conjunctive representation goes like the product of the number of nodes required to represent items and the number of nodes required to represent different times.

In principle at least, neurophysiological results can resolve the question of whether the brain contains a scale-invariant conjunctive representation of what happened when.

Memory and timing across longer time scales

In this paper we have constructed models of behavior over laboratory cognitive time scales—ranging from a few hundred milliseconds (short-term JOR) up to several hundred seconds (across-list recency and contiguity effects in free recall). There is evidence that memory manifests temporal effects over much longer scales—across days, weeks and months (Ebbinghaus, 1885/1913; Moreton & Ward, 2010; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1997). Is it reasonable to simply extend the time-line and construct models with the same form? As far as the equations are concerned, there is no limit to the range of time scales that can be addressed; the equations have the same properties for time scales arbitrarily close to zero or time scales that increase without bound. However, the time scales that can be supported in the time-line are restricted by the time constants that can be implemented in the leaky integrators $\mathbf{t}(s)$. While a reasonable argument can be made that persistently-firing neurons (Egorov, Hamam, Fransén, Hasselmo, & Alonso, 2002) can be used to construct time constants up to a few thousand seconds (Tiganj, Hasselmo, & Howard, submitted), it is hard to extend the same neurophysiological argument to longer time scales. However, even if one restricts the time-line to only a few thousand seconds it is still possible to construct behavioral models that exibit gradual forgetting over much longer time scales.

First, there are several other ways to introduce temporal effects into behavioral models based on this representation without extending the time-line. For instance, in this paper we used the match between two states of temporal history to drive behavioral effects. With the simple assumptions used in this paper, there would be no decrease in the efficacy of A as a cue for B as the retention interval following presentation of the pair increases. However, if matching was implemented in an associative memory, then allowing synaptic weights to decrease over time would cause the resulting behavioral model to show a decrease in performance with increasing retention interval. Similarly, the ability of a cue to cause the recovery of a previous state of $\mathbf{T}$ could decrease over time.

Second, environmental context changes gradually over very long time scales. Consider an experiment in which the subject studies a list of words in one session and then returns after either one week or one month for a test session. Consider the set of stimuli present in the history at the moment the subject enters the laboratory on the two sessions. Gradual changes in the environment outside of the laboratory would lead to gradual changes in the similarity of the timelines across sessions. There are likely to be correlations in the semantic (Alvarez-Lacalle et al., 2006; Anderson & Schooler, 1991; Doxas, Dennis, & Oliver, 2010) and visual stimuli (Sreekumar et al., 2014) that the subject has experienced over weeks and months. For instance, suppose that the subject listened to the news prior to the two sessions. The topics covered in the news in the first session are more likely to overlap with the topics covered in the second session if the two sessions are separated one day than if the two sessions were separated by one month. Similarly with respect to visual experience, if there was snow on the ground prior to the first visit, the probability that there was still snow on the ground prior to the second visit is higher if it was a week after the first visit rather than a month after the first visit. One can construct similar arguments for many aspects of experience (mood, songs on the radio, plotlines of TV programs, size of a child, relationship status) that are likely to contribute to the subjects' experience at the time a controlled stimulus is presented. These gradual changes in environmental circumstances could contribute directly to memory and forgetting. Similarly, recovery of episodic details of real-life situations may enable us to evaluate the age of a memory over months and even decades. For instance, if a subject recovers an episodic memory of singing their baby to sleep, the knowledge of the child's birthday would enable the subject to evaluate the age of the memory without scanning across an internal time-line extending years or decades into the past.

Conclusions

It has long been appreciated that there is a deep connection between time and memory. We described a mathematical framework that provides a mechanistic description of an ordered scale-invariant representation of internal time. We showed that this framework is sufficient to provide a concise description of fundamental behavioral results from a variety of paradigms, including an account of JORs over both short (Hacker, 1980) and longer laboratory time scales (Yntema & Trask, 1963; Hintzman, 2011), recency and contiguity effects in episodic memory (Glenberg et al., 1980; Kahana et al., 2008), and temporal mapping from conditioning (Cole et al., 1995). We reviewed neural evidence suggesting that the brain could implement something like this representation of internal time. Because of the relatively close correspondence to neural data, this representation could form the

basis of a physical model of many aspects of learning and memory.

References

Adler, A., Finkes, I., Katabi, S., Prut, Y., & Bergman, H. (2013). Encoding by synchronization in the primate striatum. *Journal of Neuroscience*, *33*(11), 4854–4866.

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.

Alvarez-Lacalle, E., Dorow, B., Eckmann, J. P., & Moses, E. (2006). Hierarchical structures induce long-range dynamical correlations in written texts. *Proceedings of the National Academy of Science, USA*, *103*(21), 7956-61.

Anderson, J., & Schooler, L. (1991). Reflections of the environment in memory. *Psychological science*, *2*(6), 396–408.

Arcediano, F., Escobar, M., & Miller, R. R. (2003). Temporal integration and temporal backward associations in human and nonhuman subjects. *Learning & Behavior*, *31*(3), 242-56.

Arcediano, F., Escobar, M., & Miller, R. R. (2005). Bidirectional associations in humans and rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(3), 301-18.

Arcediano, F., & Miller, R. R. (2002). Some constraints for models of timing: A temporal coding hypothesis perspective. *Learning and Motivation*, *33*, 105-123.

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.

Baddeley, A. D., & Hitch, G. J. (1977). Recency reexamined. In S. Dornic (Ed.), *Attention and performance VI* (p. 647-667). Hillsdale, NJ: Erlbaum.

Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neuroscience*, *32*(2), 73–78.

Barnet, R. C., Cole, R. P., & Miller, R. R. (1997). Temporal integration in second-order conditioning and sensory preconditioning. *Animal Learning & Behavior*, *25*, 221–233.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*, 173-189.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage*, *5*(1), 49-62.

Brown, G. D. A., Morin, C., & Lewandowsky, S. (2006). Evidence for time-based models of free recall. *Psychonomic Bulletin and Review*, *13*(4), 717-23.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539-76.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*(1), 127-181.

Brown, G. D. A., Vousden, J. I., & McCormack, T. (2009). Memory retrieval as temporal discrimination. *Journal of Memory and Language*, *60*(1), 194–208.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755-65.

Bunsey, M., & Eichenbaum, H. B. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, *379*(6562), 255-257.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551-581.

Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *11*(3), Suppl:327-83.

Chan, S. C., Applegate, M. C., Morton, N. W., Polyn, S. M., Chan, K. A. N. S. C., Applegate, M. C., et al. (in preparation). Recall order is predicted by category-specic neural activity of preceding items at study.

Chater, N., & Brown, G. D. A. (2008). From universal laws of cognition to specific cognitive models. *Cognitive Science*, *32*(1), 36-67.

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, *37*(1-2), 55-81.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: The MIT Press.

Cole, R. P., Barnet, R. C., & Miller, R. R. (1995). Temporal encoding in trace conditioning. *Animal Learning & Behavior*, *23*(2), 144-153.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

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.

Donkin, C., & Nosofsky, R. M. (2012). A power-law model of psychological memory strength in short- and long-term recognition. *Psychological Science*.

Doxas, I., Dennis, S., & Oliver, W. L. (2010). The dimensionality of discourse. *Proceedings of the National Academy of Science, USA*, *107*(11), 4866-71.

Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University.

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.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145-154.

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*(5), 1179–1189.

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*(2), 223-71.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*(1), 59-79.

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*, 115-135.

Farrell, S., & Lewandowsky, S. (2008). Empirical and theoretical limits on lag-recency in free recall. *Psychonomic Bulletin & Review*, *15*(6), 1236-50.

Frey, W. G., & Fozard, J. L. (1970). Effect of presentation time on the judged recency of pictures. *Journal of Experimental Psychology*, *85*(1), 105-10.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gallistel, C. R. (2007). Flawed foundations of associationism? comments on Machado and Silva (2007). *American Psychologist*, *62*(7), 682-5.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, *107*(2), 289-344.

Gershman, S. J., Schapiro, A. C., Hupbach, A., & Norman, K. A. (2013). Neural context reinstatement predicts memory misattribution. *Journal of Neuroscience*, *33*(20), 8590-5.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279-325.

Gill, P. R., Mizumori, S. J. Y., & Smith, D. M. (2011). Hippocampal episode fields develop with learning. *Hippocampus*, *21*(11), 1240-9.

Glanzer, M. (1972). Storage mechanisms in recall. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (p. 129-193). New York: Academic Press.

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.

Goldman, M. S. (2009). Memory without feedback in a neural network. *Neuron*, *61*(4), 621–634.

Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2006). An FMRI analysis of the human hippocampus: inference, context, and task awareness. *Journal of Cognitive Neuroscience*, *18*(7), 1156-73.

Grossberg, S., & Merrill, J. (1992). A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research*, *1*, 3-38.

Grossberg, S., & Merrill, J. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience*, *8*, 257-277.

Grossberg, S., & Pearson, L. R. (2008). Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: toward a unified theory of how the cerebral cortex works. *Psychological Review*, *115*(3), 677-732.

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.

Healey, M. K., & Kahana, M. J. (2013). Memory search is governed by universal principlesnot idiosyncratic strategies. *Journal of Experimental Psychology: General.*.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, *49A*, 80-115.

Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychological Review*, *77*(3), 223-233.

Hinrichs, J. V., & Buschke, H. (1968). Judgment of recency under steady-state conditions. *Journal of Experimental Psychology*, *78*(4), 574-579.

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

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition*, *38*(1), 102-15.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*, 253-271.

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. (2004). Scaling behavior in the temporal context model. *Journal of Mathematical Psychology*, *48*, 230-238.

Howard, M. W., Addis, K. A., Jing, B., & Kahana, M. (2007). Semantic structure and episodic memory. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *LSA: A road towards meaning* (p. 121-141). Mahwah, NJ: Laurence Erlbaum Associates.

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General.*, *142*(4), 1211-30.

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75-116.

Howard, M. W., Jing, B., Rao, V. A., Provyn, J. P., & Datey, A. V. (2009). Bridging the gap: Transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 391-407.

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). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269-299.

Howard, M. W., Macdonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., et al. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of Neuroscience*, *34*(13), 4692-707.

Howard, M. W., Sederberg, P. B., & Kahana, M. J. (2009). Reply to Farrell & Lewandowsky: Recency-contiguity interactions predicted by TCM. *Psychonomic Bulletin & Review*, *16*, 973-984.

Howard, M. W., Venkatadass, V., Norman, K. A., & Kahana, M. J. (2007). Associative processes in immediate recency. *Memory & Cognition*, *35*, 1700-1711.

Howard, M. W., Viskontas, I. V., Shankar, K. H., & Fried, I. (2012). A neural signature of mental time travel in the human MTL. *Hippocampus*, *22*(PMC3407826), 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.

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.

Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, *49*(4), 500-518.

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.

Jenkins, L. J., & Ranganath, C. (2010). Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *Journal of Neuroscience*, *30*(46), 15558-65.

J. I. Beare, T. (1930). On memory and reminiscence. In W. D. Ross (Ed.), *The works of aristotle* (Vol. 3). Oxford: Clarendon Press.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*, 103-109.

Kahana, M. J., Howard, M., & Polyn, S. (2008). Associative processes in episodic memory. In H. L. Roediger III (Ed.), *Cognitive psychology of memory, Vol. 2 of learning and memory - a comprehensive reference (J. Byrne, Editor)* (p. 476-490). Oxford: Elsevier.

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.

Kello, C. T., Brown, G. D. A., Ferrer-I-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., et al. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences*, *14*(5), 223-32.

Kiliç, A., Criss, A. H., & Howard, M. W. (2013). A causal contiguity effect that persists across time scales. *Journal Experimental Psychology: Learning, Memory and Cogntion*, *39*(1), 297-303.

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.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399-402.

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.

Leak, T. M., & Gibbon, J. (1995). Simultaneous timing of multiple intervals: implications of the scalar property. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*(1), 3-19.

Lewandowsky, S., Nimmo, L. M., & Brown, G. D. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language*, *58*(2), 415–428.

Lewandowsky, S., & Oberauer, K. (2009). No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1545.

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

Lewis, P. A., & Miall, R. C. (2009). The precision of temporal judgement: milliseconds, many minutes, and beyond. *Philosophical Transcripts of the Royal Society London B: Biological Sciences*, *364*(1525), 1897-905.

Lockhart, R. S. (1969). Retrieval asymmetry in the recall of adjectives and nouns. *Journal of Experimental Psychology*, *79*, 12-17.

Lohnas, L. J., & Kahana, M. J. (2014). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *40*(1), 12-24.

Luzardo, A., Ludvig, E. A., & Rivest, F. (2013). An adaptive drift-diffusion model of interval timing dynamics. *Behavioural Processes*.

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 discontiguous events. *Neuron*, *71*, 737-749.

Machado, A., & Silva, F. J. (2007). Toward a richer view of the scientific method: The role of conceptual analysis . *American Psychologist*, *62*(7), 671-681.

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.

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*(PMC2104541), 530-540.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain Research : Cognitive Brain Research*, *21*(2), 139-70.

Matzel, L. D., Held, F. P., & Miller, R. R. (1988). Information and expression of simultaneous and backward associations: Implications for contiguity theory. *Learning and Motivation*, *19*, 317-344.

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

Mensink, G.-J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*, 434-55.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation*, *1*, 359-371.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202.

Moreton, B. J., & Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin & Review*, *17*, 510-515.

Morin, C., Brown, G. D., & Lewandowsky, S. (2010). Temporal isolation effects in recognition and serial recall. *Memory & cognition*, *38*(7), 849–859.

Murdock, B. B. (1960). The distinctiveness of stimuli. *Psychological Review*, *67*, 16-31.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482-488.

Murdock, B. B. (1963). Short-term memory and paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, *2*, 320-328.

Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.

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

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological review*, *115*(3), 544.

Oberauer, K., & Lewandowsky, S. (2013). Evidence against decay in verbal working memory. *Journal of experimental psychology: General*, *142*(2), 380-411.

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*, 171-175.

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

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.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Post, E. (1930). Generalized differentiation. *Transactions of the American Mathematical Society*, *32*, 723-781.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology*, *17*, 132-138.

Primoff, E. (1938). Backward and forward associations as an organizing act in serial and in paired-associate learning. *Journal of Psychology*, *5*, 375-395.

Rakitin, B. C., Gibbon, J., Penny, T. B., Malapani, C., Hinton, S. C., & Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 15-33.

Rao, V. A., & Howard, M. W. (2008). Retrieved context and the discovery of semantic structure. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems 20* (Vol. PMC3086794, p. 1193-1200). Cambridge, MA: MIT Press.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*, 734-760.

Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes*, *44*, 147–162.

Schoenbaum, G., & Eichenbaum, H. (1995a). Information coding in the rodent prefrontal cortex. II. Ensemble activity in orbitofrontal cortex. *Journal of Neurophysiology*, *74*(2), 751-62.

Schoenbaum, G., & Eichenbaum, H. (1995b). Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *Journal of Neurophysiology*, *74*(2), 733-50.

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.

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.

Shankar, K. H., & Howard, M. W. (2012). A scale-invariant representation of time. *Neural Computation*, *24*, 134-193.

Shankar, K. H., & Howard, M. W. (2013). Optimally fuzzy scale-free memory. *Journal of Machine Learning Research*, *14*, 3753-3780.

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

Simen, P., Balci, F., Souza, L. de, Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *Journal of Neuroscience*, *31*(25), 9238-53.

Slamecka, N. J. (1976). An analysis of double-function lists. *Memory & Cognition*, *4*, 581-585.

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition*, *40*(2), 177-90.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195-231.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Science, USA*, *93*(24), 13515-13522.

Sreekumar, V., Dennis, S., Doxas, I., Zhuang, Y., & Belkin, M. (2014). The geometry and dynamics of lifelogs: discovering the organizational principles of human experience. *PLoS One*, *9*(5), e97166.

Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (submitted). A computational model for exponentially decaying persistent firing.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford.

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385-398.

Tzeng, O. J. L. (1973). Positive recency in delayed free recall. *Journal of Verbal Learning and Verbal Behavior*, *12*, 436-439.

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.

Viskontas, I. V., Knowlton, B. J., Steinmetz, P. N., & Fried, I. (2006). Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *Journal of Cognitive Neuroscience*, *18*(10), 1654-62.

Voss, R. F., & Clarke, J. (1975). 1/f noise in music and speech. *Nature*, *258*, 317–318.

Wasserman, E. A., DeLong, R. E., & Larew, M. B. (1984). Temporal order and duration: Their discrimination and retention by pigeons. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Vol. 423, p. 103-115). New York, NY: The New York Academy of Science.

Wilson, M. A., & McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, *261*, 1055-8.

Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition*, *25*, 731-739.

Yakovlev, V., Fusi, S., Berman, E., & Zohary, E. (1998). Inter-trial neuronal activity in inferior temporal cortex: a putative vehicle to generate long-term visual association. *Nature Neuroscience*, *1*(4), 310-317.

Yntema, D. B., & Trask, F. P. (1963). Recall as a search process. *Journal of Verbal Learning and Verbal Behavior*, *2*, 65-74.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, *75*(1), 168-179.

Appendix A: A simple self-terminating scanning model based on sequential access of internal time

The what-and-when representation $\mathbf{T}$ is a two-tensor with each entry indexed by a stimulus dimension and a time-point $\overset{*}{\tau}$. Fixing a single value of $\overset{*}{\tau}$ gives a vector across stimuli—a what. We assume that the subject can sequentially restrict attention to one value of $\overset{*}{\tau}$ at a time starting from the present and extending backward in time. At each step of the scan, this vector $\mathbf{T}(\overset{*}{\tau})$ is compared to the probe stimulus in the case of absolute JOR, or each of the probe stimuli in the case of relative JOR. For simplicity, we assume that there is no overlap in the representation of the probe stimuli. Recall that the values of $\overset{*}{\tau}$ present in $\mathbf{T}$ are not evenly separated, but obey Weber-Fechner spacing.

Let us denote the scaled match of a probe to $\mathbf{T}(\overset{*}{\tau})$ as $p_r(\overset{*}{\tau})$:

$$p_r(\overset{*}{\tau}) = a\langle \mathbf{f}_A | \mathbf{T}(\overset{*}{\tau})\rangle, \tag{A1}$$

where a is a free parameter, $\mathbf{f}_A$ is the vector corresponding to probe stimulus A and the $\langle\cdot|\cdot\rangle$ notation here simply reflects the inner product between two vectors. We assume that the instantaneous probability of a search for probe stimulus $\mathbf{f}_A$ terminating on $\mathbf{T}(\overset{*}{\tau})$ is given by $p_r(\overset{*}{\tau})$.

In a self-terminating scanning model, the decision terminates as soon as one query terminates. Let us first consider only one probe A. The probability that the query returns for the first time while scanning at $\overset{*}{\tau}$ is given by

$$P_f(\overset{*}{\tau}) = \left[1 - \int_0^{\overset{*}{\tau}} P_f(\overset{*}{\tau}')\ d\overset{*}{\tau}'\right] p_r(\overset{*}{\tau}) \tag{A2}$$

This has the general solution

$$P_f(\overset{*}{\tau}) = p_r(\overset{*}{\tau})g(\overset{*}{\tau})\ \exp\left[-\int_0^{\overset{*}{\tau}} p_r(\overset{*}{\tau}')g(\overset{*}{\tau}')\ d\overset{*}{\tau}'\right], \tag{A3}$$

The function $g(\overset{*}{\tau})$ is the number density of the cells along the $\overset{*}{\tau}$ axis, $g(\overset{*}{\tau}) = \frac{dN}{d\overset{*}{\tau}}$. Because the spacing of the $\overset{*}{\tau}$ axis is Weber-Fechner, the number density goes like $g(\overset{*}{\tau}) \propto |\overset{*}{\tau}|^{-1}$. Equation A2 ensures that the integral of P_f is not greater than one, allowing the integral to be understood as the probability that the query returns useful information over the course of the scan.

When a is small relative to τ_A, the time in the past at which the item was presented, then $\int_0^\infty p_r(\overset{*}{\tau})g(\overset{*}{\tau})d\overset{*}{\tau}$ is much less than one and $P_f(\overset{*}{\tau}) \simeq p_r(\overset{*}{\tau})g(\overset{*}{\tau})$. When τ_A is much less than a, then $\int_0^\infty p_r(\overset{*}{\tau})g(\overset{*}{\tau})d\overset{*}{\tau}$ is much greater than one. This means that the probability of the search returning at least some information is nearly one and more or less independent of τ_A. When τ_A is much greater than a, the probability goes down like a/τ_A. When $g(\overset{*}{\tau}) = |\overset{*}{\tau}|^{-1}$ and $\tau_A \gg a$, the mean number of rows of $\mathbf{T}(\overset{*}{\tau})$ that must be scanned for the search to return goes up with $\log(\tau_A)$.

In the absolute JOR task, we assume that the subject reports the number of the node at which the search terminates. Because of the Weber-Fechner spacing, the node number goes like $\log(|\overset{*}{\tau}|)$. In the relative JOR task, we have two queries racing to be selected. Suppose we have two probes, A and B. In this case, if we denote the probability of first return for stimulus A as $P_{f,A}(\overset{*}{\tau})$, and the probability of first return for stimulus B as $P_{f,B}(\overset{*}{\tau})$, then the probability of selecting stimulus A while querying $\overset{*}{\tau}$ will be given by

$$P_A(\overset{*}{\tau}) = P_{f,A}(\overset{*}{\tau})\left(1 - \int_0^{\overset{*}{\tau}} P_{f,B}(\overset{*}{\tau}')\ d\overset{*}{\tau}'\right) \tag{A4}$$

In order to generate predictions, all that is necessary is to pick a value for the parameter a, put explicit values of $\mathbf{T}(\overset{*}{\tau})$ given the experimentally-relevant delays τ_A and τ_B into Eq. A1 and go forward. Unfortunately, the resulting integrals cannot be calculated analytically. Accordingly, we have estimated them numerically using the R function `integrate()` in generating predicted results. The spacing between values of $\overset{*}{\tau}$ in the numerical integration was set to .005; the upper limit of integration was set to three times the longest lag.

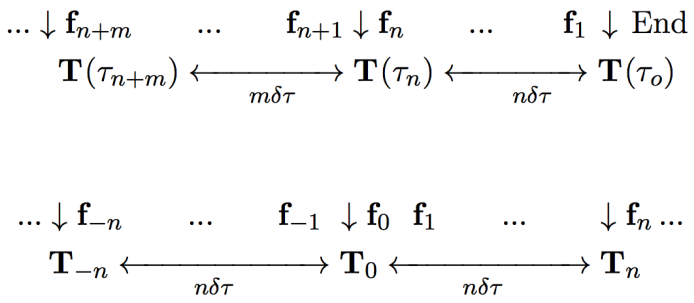

Figure 14. Schematic showing how states of **T** are numbered for the calculation of activations due to recency and contiguity. Top: Recency calculation. Bottom: Contiguity calculation.

Appendix B: Derivations of quantities used in modeling free recall

Here we derive closed form expressions for the match between two states of history at different points in a list of words. In order to make the calculation analytically simple, we assume that there is no overlap in the item representation of the words in the list and that the representation of distractors neither overlap with one another nor with the words in the list. In order to generate potential recalls, we first calculate the match between the current state of history and the state available when each word was encoded. The primary difference between our treatment of the recency effect and the contiguity effect is the state of history used to generate the match. For the recency effect, the probe history is the state of **T** when the recall test is initiated. For the contiguity effect, the probe history incorporates the effects of the just-recalled word.

Recency: Computing the match between a state of history and the encoding history of a word n steps in the past

Let's denote the list items by $\mathbf{f}_n$, where n denotes the position of the item from the end of the list. Let the n-th item be presented at time τ_n. For simplicity, we shall take the time interval between successive items to be a constant $\delta\tau$. We will further assume that there is a never-ending stream of stimuli, each presented with an interstimulus interval of $\delta\tau$. This enables us to ignore edge effects.[21] The pattern **T** at the end of the list is $\mathbf{T}(\tau_o)$ and the encoding history prior to the presentation of the item $\mathbf{f}_n$ is $\mathbf{T}(\tau_n)$. This is represented pictorially in the top of Figure 14.

At the time when an item is presented, only the columns corresponding to preceding items contain information. This means that only the predecessors of $\mathbf{f}_n$ will leave an image in $\mathbf{T}(\tau_n)$. As a consequence, only the predecessors of $\mathbf{f}_n$ contribute to the match between the $\mathbf{T}(\tau_o)$ and $\mathbf{T}(\tau_n)$. Let's denote the temporal history of an item $\mathbf{f}_{n+m}$ observed at time τ as $\mathrm{T}^m(\tau, \overset{*}{\tau})$. The contribution of the history from item $\mathbf{f}_{n+m}$ to the match between the encoding history of word n, at time τ_n and the test context, at time τ_o is then

$$A_n^m \equiv \int \mathrm{T}^m(\tau_o, \overset{*}{\tau})\, \mathrm{T}^m(\tau_n, \overset{*}{\tau})\, g(\overset{*}{\tau})\, d\overset{*}{\tau} \tag{B1}$$

[21] Because of the scale-invariance of the representation, this is a reasonable approximation.

The function $g(\overset{*}{\tau})$ is the number density of the nodes as a function of $\overset{*}{\tau}$ (see Appendix A). Now, the match between the probe context at time τ_o and the study context at time τ_n is the sum of the contribution across all of the words that preceded n:

$$p_n = \langle \mathbf{T}(\tau_n) | \mathbf{T}(\tau_o) \rangle = \sum_{m=1}^{\infty} A_n^m \tag{B2}$$

The value of the coefficients A_n^m with $g(\overset{*}{\tau}) = |\overset{*}{\tau}|^w$ can be calculated analytically, although the calculation is somewhat involved. With Weber-Fechner spacing, $w = -1$. It turns out that for any given n, the summation of A_n^ms over all values of m converges only for $w < 0$. The value of p_n when $w < 0$ is

$$p_n \simeq n^w (\delta\tau)^{(w-1)} \tag{B3}$$

The details of this calculation follow.

Detailed calculation of p_n. The calculation of A_n^m follows an analogous computation in Shankar and Howard (2012) closely (see Eqs. 5.15-5.18 in Shankar & Howard, 2012). The only difference here is that $w = -1$ with Weber-Fechner spacing rather than $w = 0$ as assumed in that paper. With $g(\overset{*}{\tau}) = |\overset{*}{\tau}|^w$, we find

$$A_n^m = \frac{k^{(1+w)}\Gamma(2k+1-w)}{k!^2} \left[\frac{(n/m)^{1-w}(1+n/m)^k}{(2+n/m)^{2k+1-w}} \right] (n\delta\tau)^{(w-1)} \tag{B4}$$

These coefficients are all finite and well defined except when both $m = n = 0$. An important property to note is that for any nonzero n, $A_n^0 = 0$, but for nonzero m, $A_0^m \neq 0$.

To compute p_n for any given n, we have to sum over the contributions from all its predecessors; that is, sum the coefficients A_n^m over all m. Note from eq. B4 that for $m \gg n$, $A_n^m \sim m^{(w-1)}$. The summation will converge for any $w < 0$ resulting in a finite p_n, while for $w >= 0$, the summation will diverge. To obtain a qualitative functional form of p_n, it is convenient to approximate the summation by an integral.

$$\begin{aligned} p_n \sim (n\delta\tau)^{(w-1)} \sum_{m=1}^{\infty} \frac{(n/m)^{1-w}(1+n/m)^k}{(2+n/m)^{2k+1-w}} &\longrightarrow (n\delta\tau)^{(w-1)} \int_0^{\infty} dm \, \frac{(n/m)^{1-w}(1+n/m)^k}{(2+n/m)^{2k+1-w}} \\ &= n^w (\delta\tau)^{(w-1)} \int_0^{\infty} dz \frac{z^{(w-1)}(1+z^{-1})^k}{(2+z^{-1})^{2k+1-w}} \end{aligned} \tag{B5}$$

In the last step, m/n is relabeled as z. This integral converges for all $w < 0$ and diverges for $w >= 0$. Thus we see that for $w < 0$, $p_n \sim n^w (\delta\tau)^{(w-1)}$.

The effect of delay intervals on the PFR. Let the number of distractor stimuli between two words be $D-1$, making the total time elapsed between the presentation of two successive words (including the presentation of the first word) just D. Let us denote the duration of the distractor interval after the last word in the list and prior to the recall test as d. Now the positions of the list items in the sequence of stimuli are $(d + nD)$, where n is the word number starting from the end of list. For the PFR, Eq. 2 simplifies as

$$\mathrm{PFR}(n) = \frac{(d+nD)^{-b}}{\sum_m (d+mD)^{-b}}, \tag{B6}$$

where the summation over m goes over all ℓ words in the list.

In order to fit this expression to experimental data, we simply set d and D to correspond to the experimental values. In continuous distractor experiments, the delay between items is not repeated at the end of the list. To get the delays in Eq. B6 to correspond to the experimental values, we let n in Eq. B6 run from 0, corresponding to the last word in the list, to $\ell-1$, corresponding to the first word in the list in continuous distractor experiments. In final free recall experiments, the delay after presentation of an item (consisting primarily of the recall period and the time between lists) is repeated after the last item. To get Eq. B6 to correspond to the delays in the actual experiments, we let n run from 1, corresponding to the last list in the session.

To describe final free recall, ideally we would sum over the contributions from all items within each list and compare the relative activations of different lists. However for the purpose of a qualitative demonstration, we used an approximation. If all of the ℓ words in a list had precisely the same activation, then to calculate the relative probability of recalling a word from the list it is sufficient to treat the entire list as if it is a single item. To estimate the probability of initiating final free recall with a word from list n, we use Eq. B6 with d and D chosen in accordance with the experiment.

Contiguity calculations

In order to calculate the contiguity effect caused by the recalled item in isolation from the recency effect, let us now assume that a lot of time has elapsed since the end of the study list and that one item from the center of the list has just been recalled. Our aim is to predict the probability that the other list items would be recalled next as a function of the distance from the just-recalled word to the potential recalls. Let the recalled item be labeled as 0, and let us relabel the neighboring study list items by successive integers (Figure 14, bottom). It will be necessary to keep track of the activation for items that followed $\mathbf{f}_o$, namely $\mathbf{f}_{+1}$, $\mathbf{f}_{+2}$, and so on, separately from the activation of the preceded $\mathbf{f}_o$, namely $\mathbf{f}_{-1}$, $\mathbf{f}_{-2}$, etc..

Suppose that $\mathbf{f}_o$, the stimulus initially presented at time τ_o is repeated (and remembered) at τ_r. In order to accommodate recovery of the temporal history, we assume that when $\mathbf{f}_o$ is repeated, we momentarily alter the differential equation controlling $\mathbf{t}(s)$ as follows:

$$\frac{d\mathbf{t}(s)}{d\tau} = -s\mathbf{t}(s) + \gamma\mathbf{t}(\tau_o, s) + (1-\gamma)\,\mathbf{f}(\tau_o) \tag{B7}$$

Ordinarily, $\mathbf{t}$ evolves from moment-to-moment with $\gamma = 0$ (Shankar & Howard, 2012). The factor γ must be less than one and controls the strength of reinstatement of the previous temporal history. When this new state of $\mathbf{t}(s)$ is operated on by $\mathbf{L}_{\mathrm{k}}^{-1}$, this gives rise to two new components of $\mathbf{T}$. One is the repetition of the stimulus itself, weighted by $1-\gamma$ that enters at $\overset{*}{\tau} = 0$ and triggers the same sequence of activations in $\mathbf{T}$ it did when the item was initially presented (except weighted by $1-\gamma$). The other is $\gamma\ \mathbf{L}_{\mathrm{k}}^{-1}\mathbf{t}(\tau_o)$. This is the state $\mathbf{T}(\tau_o)$ that was present just as $\mathbf{f}_o$ was originally presented. Note that immediately after τ_r, the former component is not an effective cue for recall of any other list item. That is because no item in the list was experienced with the recalled item 0 s in the past. $\mathbf{f}_{+1}$ was experienced with the recalled item $\delta\tau$ seconds in the past. On the other hand, $|\mathbf{T}_0\rangle$ is an effective cue for other list items immediately after recovery. The match to the encoding

history for an item in the forward direction due to $|\mathbf{T}_0\rangle$ is given by $p_n = \gamma\langle\mathbf{T}_n|\mathbf{T}_0\rangle$, and in the backward direction is given by $p_{-n} = \gamma\langle\mathbf{T}_{-n}|\mathbf{T}_0\rangle$. Since it is the same amount of time separation and item separation between $\mathbf{T}_{-n}$ and $\mathbf{T}_0$ as is between $\mathbf{T}_0$ and $\mathbf{T}_n$, the prediction generated should be symmetric, meaning p_n and p_{-n} are equal and are given by Eq. B2. Hence immediately after the reinstatement of $|\mathbf{T}_0\rangle$, the CRP should be symmetric with

$$p_{-n} = p_n = \gamma\ \langle\mathbf{T}_n|\mathbf{T}_0\rangle = \gamma \sum_{m=1}^{\infty} A_n^m \tag{B8}$$

If we wait for some time following the context reinstatement, the lag-CRP will become asymmetric favoring the forward direction. This happens because both $\mathbf{T}_0$ and the item representation entering at $\overset{*}{\tau} = 0$ evolve in time. First let's consider the effect on $\mathbf{T}_0$. Consider the state obtained by evolving $\mathbf{T}_0$ for a time $\tau = r\delta\tau$. Let us denote it by $\mathbf{T}_{0+r}$. Then the match for the n-th item in the forward and backward direction can be deduced to be

$$p_{-n} = \gamma\ \langle\mathbf{T}_{-n}|\mathbf{T}_{0+r}\rangle \quad = \quad \gamma \sum_{m=1}^{\infty} A_{n+r}^m \qquad \text{and} \tag{B9}$$

$$p_n = \gamma\ \langle\mathbf{T}_n|\mathbf{T}_{0+r}\rangle \quad = \quad \gamma \sum_{m=1+\min(n,r)}^{\infty} A_{|n-r|}^m \tag{B10}$$

In addition, the pattern of activity in the $\mathbf{T}$ column corresponding to the recalled item $\mathbf{f}_0$ will generate further asymmetry in the forward direction because the activity in this $\mathbf{T}$ column only matches the encoding history of items in the forward direction. The match to item $\mathbf{f}_n$ due to this activity is simply A_{n-r}^r when $n > r$ and is A_{r-n}^n when $r > n$. When $r = 0$, this quantity is zero, hence immediately after the recall the CRP would still be symmetric and the asymmetry would grow with time. Combining the prediction from this part along with the prediction from the reinstated context, we have

$$p_{-n} \quad = \quad \gamma \sum_{m=1}^{\infty} A_{n+r}^m \qquad \text{and} \tag{B11}$$

$$p_n \quad = \quad \gamma \sum_{m=1+\mu}^{\infty} A_{|n-r|}^m + (1-\gamma) A_{|n-r|}^{\mu} \tag{B12}$$

where μ=min(n, r), the minimum of the numbers n and r.