The Interaction between Semantic Representation and Episodic Memory

Jing Fang
jing.fang@rub.de
Mercator Research Group "Structure of Memory," Institute for Neural Computation,
and Faculty of Psychology, Ruhr University Bochum, Bochum 44801, Germany

Naima Rüther
die_ima@yahoo.de
Faculty of Psychology, Ruhr University Bochum, Bochum 44801, Germany

Christian Bellebaum
christian.bellebaum@hhu.de
Institute of Experimental Psychology, Heinrich Heine University Düsseldorf,
Düsseldorf 40225, Germany

Laurenz Wiskott
laurenz.wiskott@ini.rub.de
Institute for Neural Computation, Ruhr University Bochum,
Bochum 44801, Germany

Sen Cheng
sen.cheng@rub.de
Mercator Research Group "Structure of Memory" and Institute for Neural
Computation, Ruhr University Bochum, Bochum 44801, Germany

The experimental evidence on the interrelation between episodic memory and semantic memory is inconclusive. Are they independent systems, different aspects of a single system, or separate but strongly interacting systems? Here, we propose a computational role for the interaction between the semantic and episodic systems that might help resolve this debate. We hypothesize that episodic memories are represented as sequences of activation patterns. These patterns are the output of a semantic representational network that compresses the high-dimensional sensory input. We show quantitatively that the accuracy of episodic memory crucially depends on the quality of the semantic representation. We compare two types of semantic representations: appropriate representations, which means that the representation is used to store input sequences that are of the same type as those that it was trained on, and inappropriate representations, which means that stored inputs differ

from the training data. Retrieval accuracy is higher for appropriate representations because the encoded sequences are less divergent than those encoded with inappropriate representations. Consistent with our model prediction, we found that human subjects remember some aspects of episodes significantly more accurately if they had previously been familiarized with the objects occurring in the episode, as compared to episodes involving unfamiliar objects. We thus conclude that the interaction with the semantic system plays an important role for episodic memory.

1 Introduction

When Tulving (1972) introduced the distinction between episodic memory and semantic memory, he conceived of the former as information about specific personally experienced events and the latter as general knowledge about the world divested of a specific spatiotemporal context. An example for episodic memory would be to remember having seen a butterfly in the morning in the garden. A semantic memory would be knowing what the word *butterfly* means. Tulving (1972) also described the interaction between the episodic and semantic systems, which he later formalized as the SPI model (Tulving, 1995). The model involves three hierarchically arranged components: the perceptual, semantic, and episodic systems (see Figure 1A). Information flow among these components is process specific: serial encoding, parallel storage, and independent (SPI) retrieval (Tulving, 1995; Tulving & Markowitsch, 1998). Serial encoding means that incoming information must first pass from the perceptual to the semantic system and is then encoded as episodic memory. As a consequence, acquisition of new episodic memory is affected by information in semantic memory. The output of a given system can be transmitted to the next level or stored at that level, or both. Hence, parallel storage implies that different aspects of the incoming information are stored in different systems. Finally, the stored information can be retrieved independently from each system (Tulving, 2002). Since Tulving's early work, a vast number of studies have investigated the relationship between episodic and semantic memory, as well as their neural underpinnings. The findings broadly fall into one of three categories: (1) episodic and semantic memory are two separate memory systems, (2) they are part of a single memory system, and (3) they are separate but strongly interacting systems.

In the first category, studies focus on the dissociation between the two forms of memory. Neuropsychological studies show that patients with medial temporal lobe damage have a severe impairment in episodic memory (Bayley, Hopkins, & Squire, 2006; Rosenbaum et al., 2008), while their semantic memory is largely spared (Manns, Hopkins, & Squire, 2003). Conversely, patients with semantic dementia have relatively spared episodic memory (Chan et al., 2001; Graham & Hodges, 1997). These findings
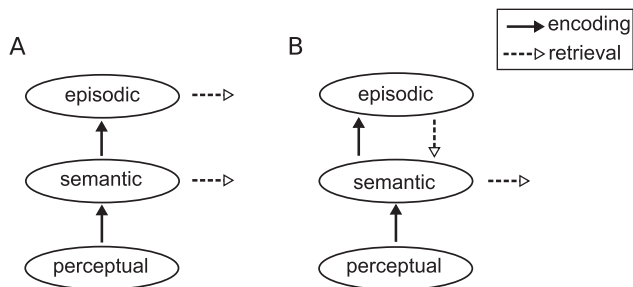
Figure 1: Schematic illustration of the relationship between memory systems. (A) Serial encoding, parallel storage, and independent (SPI) retrieval model suggested by Tulving (1995). (B) The model proposed here. In contrast to the SPI model, information is retrieved from the episodic system through the semantic system, that is, retrieval is not independent. In addition, our model makes different assumptions about the nature of the components. For more details, see the main text.

suggest a double-dissociation of the two memory systems. This dissociation has also been observed in neuroimaging studies, where tasks thought to engage the different memory systems engaged distinct sets of brain regions (Wiggs, Weisberg, & Martin, 1998; Graham, Kropelnicki, Goldman, & Hodges, 2003; Düzel, Habib, Guderian, & Heinze, 2004; Gilboa, 2004; Maguire, 2001). In general, episodic memory is thought to crucially rely on medial temporal lobe structures, in particular, the hippocampus, whereas semantic memory is processed primarily in the neocortex (Svoboda, McKinnon, & Levine, 2006; Huth, Nishimoto, Vu, & Gallant, 2012). Nevertheless, it remains contentious how episodic memory could be conceptually distinguished from semantic memory (Cheng & Werning, 2016; Tulving, 1985; Conway, 2009; Clayton, Bussey, & Dickinson, 2003; Klein, 2013b; Suddendorf & Corballis, 1997).

In the second category, numerous studies reported commonalities in neural activations across tasks involving semantic and episodic memories (Rajah & McIntosh, 2005; Burianova, McIntosh, & Grady, 2010; Binder, Desai, Graves, & Conant, 2009; Ryan, Cox, Hayes, & Nadel, 2008). The unitary system view therefore proposes that a single declarative memory system subserves both episodic and semantic memory (Rajah & McIntosh, 2005; Baddeley, 1984; Burianova & Grady, 2007). Memory encoding is thought to always be contextual (Baddeley, 1984), but at retrieval, memories may, or may not, become decontextualized depending on the task demands (Rajah & McIntosh, 2005; Westmacott & Moscovitch, 2003; Yassa & Reagh, 2013). In the most extreme version of this view, the same memory trace could be retrieved as either episodic or semantic memory, depending on whether the

retrieval was associated with autonoetic or noetic consciousness, respectively (Klein, 2013a).

Finally, the intermediate view between the first two is that episodic and semantic memory constitute two separate memory systems but interact strongly with each other (Nyberg, Forkstam, Petersson, Cabeza, & Ingvar, 2002; Cabeza & Nyberg, 2000; Gabrieli, Poldrack, & Desmond, 1998; Greenberg & Verfaellie, 2010; Cheng, Werning, & Suddendorf, 2016). Many experimental studies suggest that semantic memory affects episodic encoding (Ween, Verfaellie, & Alexander, 1996; Graham, Simons, Pratt, Patterson, & Hodges, 2000; Kinsbourne, Rufo, Gamzu, Palmer, & Berliner, 1991; Maguire, Kumaran, Hassabis, & Kopelman, 2010) and retrieval (Martin & Chao, 2001; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001; Mion et al., 2010; Spreng, Mar, & Kim, 2009). Another line of evidence shows that episodic memory is improved when the to-be-remembered information is consistent with prior semantic knowledge (Anderson & Pichert, 1978; Bartlett & Kintsch, 1932; Kan, Alexander, & Verfaellie, 2009; Hemmer & Steyvers, 2009). The levels-of-processing phenomenon also supports the special role of semantic information in episodic memory (Craik & Lockhart, 1972; Craik & Tulving, 1975). Furthermore, semantic information has even been suggested to contribute to mental time travel in the future (Irish, Addis, Hodges, & Piguet, 2012; Duval et al., 2012), which is closely linked to episodic memory (Suddendorf & Corballis, 1997). On the other hand, episodic memory has been suggested to affect both the formation and retrieval of semantic memory (Verfaellie, 2000; Kitchener, Hodges, & McCarthy, 1998; Westmacott, Black, Freedman, & Moscovitch, 2004; Kopelman, Stanhope, & Kingsley, 1999).

These disparate views cannot be easily reconciled with each other or with the SPI model, and so the relationship between the episodic and semantic systems remains controversial. Here, we suggest studying the computational function of the interaction between the episodic and semantic systems as a way forward in this controversy. We develop an abstract computational framework for the interaction between the episodic and semantic systems, which derives, but also significantly differs, from Tulving's SPI model. We find in our model that the quality of the semantic representation is important for the accuracy of episodic memory retrieval. We test this prediction in a behavioral experiment and find that subjects indeed remember some key aspects of episodic memory more precisely when their semantic representation was better tuned to the objects in the episode.

2 Model of the Interaction between the Semantic and Episodic Systems

From Tulving's SPI model, we adopt the component structure and the process-specific operation of the combined systems, but make very different assumptions about the nature of the components and the operation of the system (see Figure 1B). First, the SPI model does not specify the formats in which the episodic and semantic systems encode and store information. In

our model, these formats critically determine how the episodic and seman-
tic systems function. We have previously suggested that episodic memo-
ries are best represented as sequences of neural activity patterns (Cheng,
2013; Cheng & Werning, 2016; Cheng et al., 2016). More specifically, the
recurrent CA3 network encodes sequences during the experience and re-
plays them later during retrieval (Levy, 1996; Lisman, 1999; Buhry, Azizi,
& Cheng, 2011; Cheng, 2013). By contrast, semantic memory in our model
is represented by static neural activity patterns. As in the SPI model, en-
coding in our model is serial. The perceptual system sends inputs to the
semantic system, which generates a semantic representation of the inputs.
These semantic representations have been learned previously on a body of
input data. During the experience of an episode, sequences of semantic pat-
terns are stored by the episodic system as episodic memory. Specifically, the
sequence is stored by associating each element of the sequence with its suc-
ceeding element. In summary, episodic memory is defined as a sequence
of semantic representations. This feature makes our model different from
other theoretical and computational models of the semantic and episodic
systems.

Second, in contrast to the SPI model, stored information cannot be re-
trieved independently in our model. Instead, episodic memories have to
be retrieved through the semantic system (see Figure 1B), that is, the infor-
mation stored in the episodic system must first pass through the semantic
system before it can be retrieved (Cheng et al., 2016). This view is consistent
with empirical observations suggesting that the semantic system is impor-
tant for episodic retrieval (Gilboa, 2004; Mion et al., 2010; Spreng et al., 2009;
Martin & Chao, 2001; Mangels, 1997; Stuss, Craik, Sayer, Franchi, & Alexan-
der, 1996). However, in our current computational model, we do not imple-
ment this interaction during retrieval explicitly. To study the properties of
stored episodic sequences, they are directly retrieved from the episodic sys-
tem. The retrieval of an episodic sequence is initiated by providing a single
pattern as a retrieval cue and then proceeds from one element to the next.

In the following, we describe the elements of our model in more detail.

2.1 Input Patterns. As input patterns, we used sequences of images
containing one of six letter-like objects: T, L, U, E, H, and X (see Figure 2).
Since we wanted to train the model of the semantic system on sequences of
a large number of images, we had to limit the images to 30×30 gray-scale
pixels to keep the memory requirement and run-time to manageable levels.
However, because of the severe pixelation at this resolution, we first gen-
erated 300×300-pixel black-and-white images and then scaled them down
to 30×30 gray-scale images by averaging across 10×10 patches. Image se-
quences show one object rotating while the object's center (r_1, r_2) describes
a Lissajous curve,
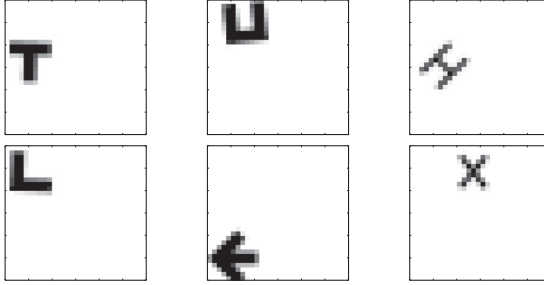
$$r_i(t) = a_i \sin (b_i t + c_i), \tag{2.1}$$

Figure 2: Examples of the images used in the modeling study. Images show one of six possible objects, which we refer to as T, U, H, L, E, and X.

where $i = 1, 2$ and $a_1 = a_2 = 100$ due to the square input space. The ratio $\frac{b_1}{b_2} = \frac{\pi}{3}$ was set to an irrational number so that theoretically, the trajectory would never repeat. However, in practice, the trajectory repeats due to the discretization of floating-point numbers on the computer and the images' pixelation. The rotation was described by

$$\varphi(t) = \omega t + \varphi_0, \tag{2.2}$$

where φ is the orientation and $\omega = 0.025e$ (e is Euler's number, again irrational).

We also used a random walk trajectory in the testing phase, where the object moves horizontally and vertically in each time step. The steps in the two directions are drawn independently from a normal distribution $v \sim N(5, 2.2)$ to reduce repetitions in the position. If a step would take the object beyond the boundary, the object is reflected on the boundary instead. The rotation of the object also follows a random walk, where the steps are drawn from $\delta_\varphi \sim N(0, (0.035e)^2)$.

The parameters of the object's movement were chosen to ensure that semantic representations could be learned from the input data in a systematic way. Note that the semantic representation does not learn the movement statistics itself, so that in our model, the movement statistics is not important for the model after the semantic system has been trained.

2.2 Model of the Semantic System

2.2.1 Slow Feature Analysis. In this study, we focus on low-dimensional (semantic) representations of high-dimensional sensory inputs and neglect other aspects that have previously been ascribed to the semantic systems, such as semantic memory, concepts, and schemas. We chose the slow feature analysis (SFA) algorithm as the model for semantic learning because

it was developed to extract slowly varying features that are hidden in more quickly varying data (Wiskott & Sejnowski, 2002). It is guided by the observation that meaningful information contained in an input stream varies more slowly than the individual components of the input. For instance, the pixel values in an image vary quite quickly as an object rotates in a fixed position, whereas the position and identity of the object do not change. Although SFA learns on changes in time, once trained, the network can extract the semantic representation from an isolated pattern of input, consistent with our model's assumption that semantic memory is represented by single patterns of neural activity. SFA has been shown to account for experimentally observed receptive field properties of complex cells in the visual system (Berkes & Wiskott, 2005) and oriospatial cells including place, grid and head direction cells in the hippocampal formation (Franzius, Sprekeler, & Wiskott, 2007). At a technical level, SFA finds instantaneous scalar input-output functions that generate slowly varying output signals from a quickly varying input signal. In a given function space F and given a multidimensional input signal $\mathbf{x}(t)$, SFA finds a set of functions $\{g^{(1)}(\mathbf{x}), g^{(2)}(\mathbf{x}), \ldots, g^{(i)}(\mathbf{x}), \ldots\}$, where $g^{(i)}(\mathbf{x}) \in F$, such that the output signals $\{y^{(1)}(t), y^{(2)}(t), \ldots, y^{(i)}(t), \ldots\}$, where $y^{(i)}(t) := g^{(i)}(\mathbf{x}(t))$, minimize

$$\Delta(y^{(i)}) := \left\langle (\dot{y}^{(i)})^2 \right\rangle_t \tag{2.3}$$

under the following constraints:

$$\left\langle y^{(i)} \right\rangle_t = 0 \quad \text{(zero mean)}, \tag{2.4}$$

$$\left\langle (y^{(i)})^2 \right\rangle = 1 \quad \text{(unit variance)}, \tag{2.5}$$

$$\forall j \text{ s.t. } j < i : \left\langle y^{(j)} y^{(i)} \right\rangle_t = 0 \quad \text{(decorrelation and order)}. \tag{2.6}$$

Equation 2.3 introduces the Δ-value, which is a measure of the slowness of the signal $y^{(i)}(t)$. Equations 2.4 and 2.5 avoid the trivial solution of a constant function, for which $\Delta = 0$. The constraint in equation 2.6 ensures that SFA does not yield the same feature twice and induces an order such that $y^{(1)}(t)$ is the slowest-varying output and $\Delta y^{(j)} < \Delta y^{(i)}$ for $j < i$. The ordering is achieved because the optimization is applied successively for increasing i and the components with larger index i have more constraints placed on them. The first component ($i = 1$) has no constraint from equation 2.6 since there is no $j < i$. Hence, $y^{(1)}$ will be the optimal solution, the slowest varying feature. The larger the index i, the more components j satisfy the condition $j < i$, and thus the more constraints have to be satisfied by $y^{(i)}$. As a result, the solutions become less optimal (i.e., the features vary faster), for larger i. The learned functions $g^{(i)}$ operate on single input patterns, which makes SFA fundamentally different from low-pass filtering.

2.2.2 Hierarchical Network Structure and Training. We used a hierarchical SFA network to model the semantic system. Information is first extracted locally and then integrated into more and more global and abstract features at each level of processing (see Figure 3A). The network consists of converging layers of SFA nodes. In the first layer consisting of 6×6, each node receives inputs from a 15×15 pixel region in the image space, which is called the receptive field. In each direction, the receptive fields of two neighboring nodes overlap by 12 pixels. In the second layer, each one of 2×2 nodes receives inputs from a 4×4 patch of first-layer nodes with an overlap of two nodes. These nodes converge onto a single SFA node in the third layer. In this setup, all nodes in a given layer together cover the full image space. The activity of the node in the top layer is taken to be the output of the semantic system in our memory model.

In each SFA node, the same set of processing steps is performed (see Figure 3A, top right). The first linear SFA stage performs dimensionality reduction, whereby the input dimensionality is reduced to 48. The following quadratic expansion allows for nonlinear features and increases the dimensionality to 1224. A final linear SFA stage reduces the dimensionality to 32, except for the top layer, where the output dimensions are reduced to 4. Since each layer realizes a polynomial of degree 2, the network as a whole computes subsets of polynomials of degree $2^3 = 8$.

The hierarchical network was implemented in Python using the MDP library (Zito, Wilbert, Wiskott, & Berkes, 2008). It was trained sequentially from bottom to top on sequences of 10,000 images in each training session; longer data sets did not improve the training. With our choice of the object's movement statistics, we ensure that the slowest features that emerged from the SFA network correspond to the coordinates of the object's center and its orientation (see e.g., Figures 3B and 4). The four slowest features were considered the semantic representation, denoted by $\mathbf{y}_t$ on the input at time point t. Note that SFA learns to extract a more abstract representation of a single input image; it does not learn the movement statistics.

2.3 Episodic Sequence Storage and Retrieval. After the sensory input is processed by the semantic system, sequences of semantic representations are stored in the episodic system. To be certain about what computations occur in the sequence network, we chose to use a highly simplified algorithmic model for sequence storage and retrieval. In our model, the episodic sequences are stored element by element (i.e., each element is stored individually). To preserve the sequential information, each element $\mathbf{y}_t$ is associated with the next element in the sequence $\mathbf{y}^*_{t+1}$, which we refer to as the associated cue (see Figure 3C). It serves as a retrieval cue when the sequence has to be retrieved. Hence, the information about the temporal ordering of the elements in the sequence is not available at a global level, only on a pairwise basis.
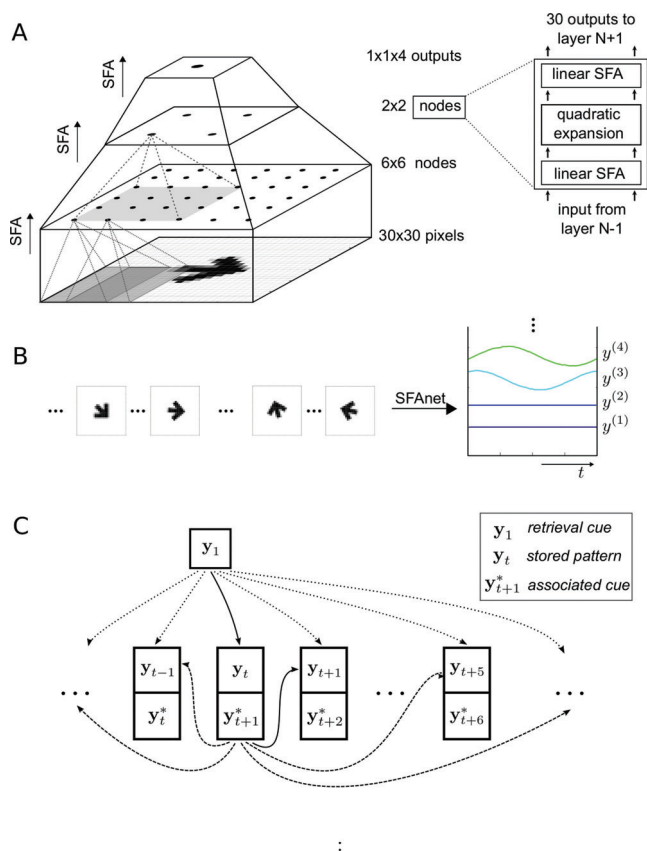
Figure 3: Illustration of the algorithmic model of the semantic system and sequence storage network. (A) Hierarchical network of slow feature analysis (SFA) as a model of the semantic system. The dots in each layer symbolize SFA nodes. The gray patches indicate the receptive field of each node, partially overlapping with the neighboring nodes' receptive fields. Each node performs a number of processing steps, as visualized on the right-hand side and detailed in the text. (B) Example of an image sequence and the corresponding output of the top node. The four curves indicate the four features ordered by their slowness (from bottom to top). They represent the two coordinates of the object's center and the sine and cosine of the object's orientation. (C) The sequence storage network as a model of the episodic system. Each element in a sequence $\mathbf{y}_t$ is isolated from the rest of the sequence and stored in the network. The sequence information is preserved by associating each element with the subsequent element in the sequence, denoted by $\mathbf{y}_{t+1}^*$ (association indicated by two fused squares). The arrows illustrate the retrieval process. The first sequence element is provided externally to initiate retrieval. At each retrieval step, the retrieval cue is compared to all stored patterns, and the closest one is retrieved (solid line). At each successive step, the associated cue becomes the retrieval cue to the next element.
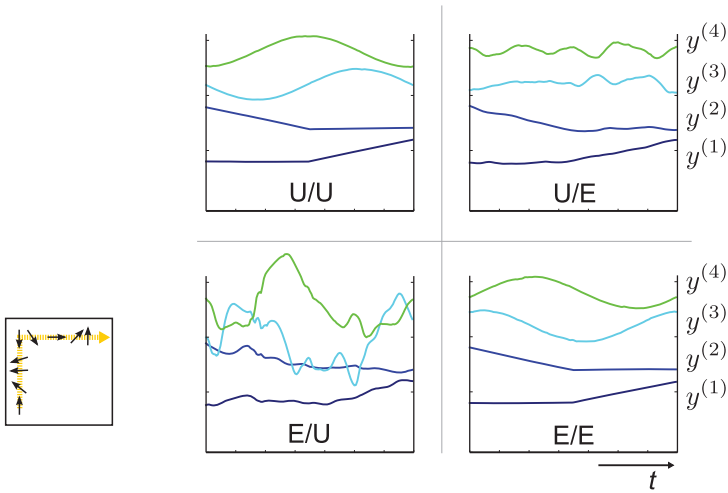
Figure 4: Example of four SFA features for the same movement trajectory when using appropriate and inappropriate representations. The panel in the lower left shows the trajectory (yellow lines with arrow) and orientation (black arrows) of the object. The four other panels show the values of the four slowest features for four different combinations of objects and networks. The notation A/B indicates that the SFA network was trained on A and tested on B. Within the four panels, the features are sorted according to their slowness, with the slowest at the bottom. These features correspond to the x- and y-coordinate of the object, and the sine and cosine of the object's orientation. While the appropriate representations (U/U and E/E) yield smoother features, the inappropriate representations (U/E and E/U) represent the features to some extent as well, particularly the x- and y-coordinates, even though they were not trained on that particular input.

This model is sufficient for our purposes here, since we focus on the interaction between the semantic and the episodic system, not on sequence storage itself. We believe that this algorithm approximates the effective computation of heteroassociations between one pattern in a sequence and the next in area CA3 in the hippocampus (Jensen & Lisman, 1996; Lisman, 1999). Our model is akin to the chaining model, which appears in various theories of memory (Ebbinghaus, Ruger, & Bussenius, 1913; Jones, Beaman, & Macken, 1996; Kieras, Meyer, Mueller, & Seymour, 1998; Lewandowsky & Murdock, 1989; Murdock, 1993; Wickelgren, 1965). It assumes that during storage, associations are formed between successive items. During retrieval, each item acts as the retrieval cue for the subsequent item (Ebbinghaus et al., 1913; Lewandowsky & Murdock, 1989; Murdock, 1993; Lashley, 1951). While the chaining model does not account for some features of

serial-order memory (see section 5), it appears to be appropriate for modeling episodic memory for our purposes.

The retrieval process in our model is triggered by providing an initial retrieval cue. We always use the first pattern of a sequence $\mathbf{y}_1$. The algorithm searches among all stored patterns for the pattern $\mathbf{y}'_1$ that has the smallest Euclidean distance to the cue. This process bears some resemblance to global matching models of recognition memory (Hintzman, 1984; Murdock, 1993; Logan, 2002; Shiffrin & Steyvers, 1997). The retrieval step is iterated for the subsequent elements in the sequence, where the pattern $\mathbf{y}^*_{t+1}$ associated with the just retrieved pattern $\mathbf{y}'_t$ serves as the new cue. This process continues for a certain number of retrieval steps. The process described up to this point is noise free and would retrieve the sequences perfectly as long as no identical patterns were stored in the sequence network. However, biological neural networks have to operate in the face of internal and external noise. Therefore, we add a noise term $\epsilon^{(k)}_{t+1} \sim N(0, \sigma^2)$ during each retrieval step,

$$s' = \operatorname*{argmin}_s \left\| \mathbf{y}^*_{t+1} + \boldsymbol{\epsilon}_{t+1} - \mathbf{y}_s \right\|, \tag{2.7}$$

$$\mathbf{y}'_{t+1} = \mathbf{y}_{s'}, \tag{2.8}$$

where $\mathbf{y}_s$ stands for the stored patterns and s' stands for the index of the retrieved pattern. Since we found that the retrieval noise plays an important role in the retrieval process, we present many results as a function of its standard deviation σ.

The retrieval error is quantified by the Euclidean distance between the original and the retrieved sequence elements;

$$e(t) = \left\| \mathbf{y}'_t - \mathbf{y}_t \right\|. \tag{2.9}$$

2.4 Study Design. We used two categories of objects, say A and B. For each category, we used sequences of images of the objects to train a semantic network. We then used the semantic network trained on category A to store 30 sequences, each consisting of 50 elements (unless noted otherwise), drawn from A. This case represents sequence storage using an appropriate semantic representation and is denoted by A/A. We then tested episodic memory performance by trying to retrieve the stored sequences and calculating the retrieval error. Memory performance in the A/A case was compared to the case where the same semantic network was used to store sequences of inputs drawn from the other category B, denoted by A/B. In this case, the semantic representation is inappropriate. To ensure that any difference in memory retrieval between appropriate and inappropriate cases does not merely reflect an asymmetry between categories A and B, we performed the same comparison where the roles of categories A and B were swapped (i.e., B/B versus B/A). Note that we compared the retrieval

performance on the basis of the same representation but not the same test object. This is consistent with the intrasubject design in our behavioral experiment (see section 4.1), in which each subject has a given representational system but is tested with different objects.

An example for the outputs of two appropriate and two inappropriate representations is shown in Figure 4. Two SFA networks were trained on U and E, respectively. Both networks extracted the location and orientation of the trained object in the same order of slowness (see Figure 4, U/U and E/E). Although each network was trained on only one of the objects, they nevertheless also represent features of other objects to some extent (U/E and E/U). Put differently, the learned semantic representations generalize to other objects to a limited degree.

2.4.1 Whitening. Inspection of example outputs from appropriate and inappropriate semantic networks revealed that the variance of the output differs greatly between the two cases (see Figure 5). This difference in variance leads to a bias when comparing retrieval errors. The larger the variance, the larger the absolute retrieval errors. In the (unlikely) extreme case, where all inputs are mapped onto a single output, the variance is zero and all retrieval errors would vanish. To avoid biases due to different variances, we normalized the data by whitening; the data are transformed such that the covariance matrix is the identity matrix (see Figure 6). To determine the transformation, we generated a small sample of input patterns (3000) and fed them through the trained semantic network to obtain the data matrix $Y = [\mathbf{y}_1, \ldots, \mathbf{y}_{3000}]$. We then determined the transformation matrix W such that the new data matrix,

$$\tilde{Y} = WY, \tag{2.10}$$

has unit covariance. This occurs if $W = D^{-1/2}V^T$, where D is a diagonal matrix whose elements are the eigenvalues of the covariance matrix of Y and V is the matrix of corresponding eigenvectors. By using the transformed data to calculate the retrieval error, we effectively calculated the relative error $e(t) = \left\| \tilde{\mathbf{y}}'_t - \tilde{\mathbf{y}}_t \right\|$ instead of the absolute error. For simplicity, we omitted the tilde symbol even when we used the transformed data.

3 Semantic Representations Affect the Quality of Episodic Memory ___

3.1 Inappropriate Semantic Representation Yields Larger Episodic Retrieval Errors. The main result of this study is that a retrieved sequence more closely matches the stored sequence when an appropriate representation is used, as compared to when an inappropriate representation is used (see Figure 7). More quantitatively, the episodic retrieval error was consistently lower for appropriate representations (see Figure 8A, blue curves)
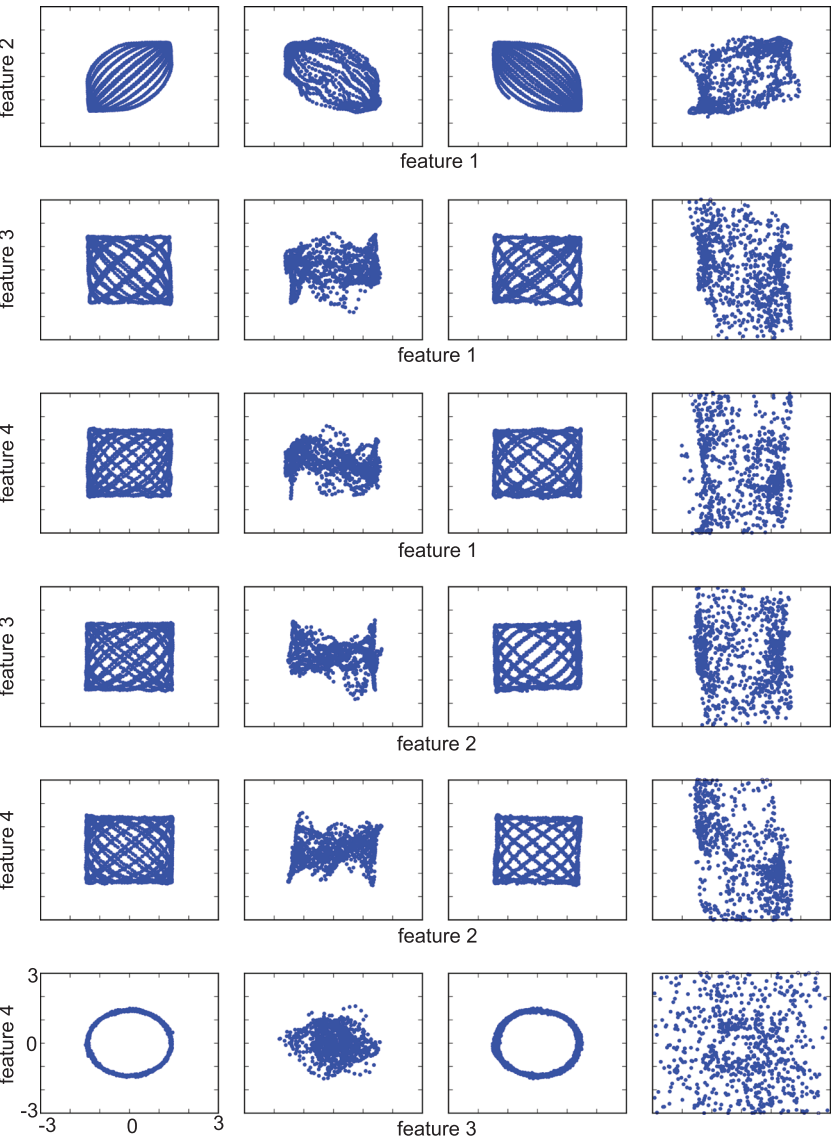
Figure 5: Example of 3000 stored patterns. The four columns show data for U/U, U/E, E/E, and E/U, respectively. The 4D pattern vectors are visualized by plotting each dimension against another. Note that the temporal structure of the sequences is not visualized.
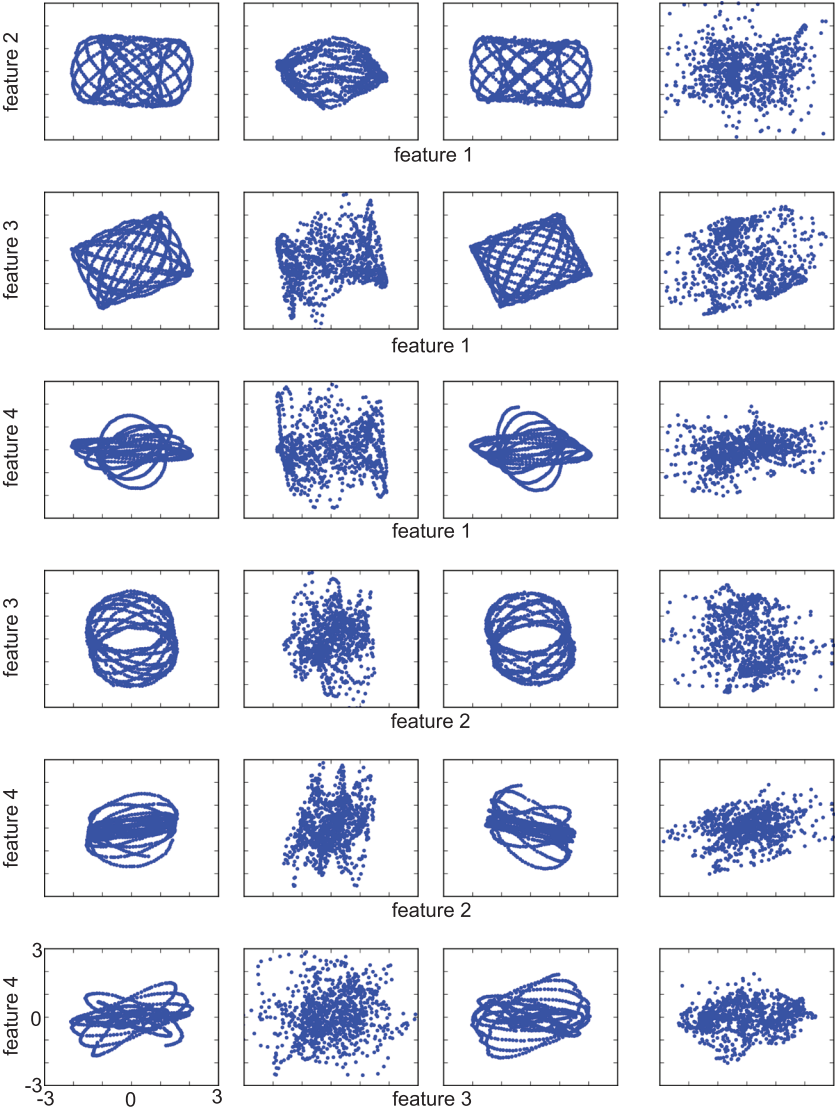
Figure 6: Same data as Figure 5 after the whitening transformation.

than for inappropriate representations (see Figure 8A, red curves). In the remainder of this letter, we expand on this result, elucidate the underlying mechanisms, and test our model's predictions experimentally.

The level of retrieval noise has an important effect on episodic retrieval (see Figure 8A). Without it, episodic retrieval in all four cases is flawless.

Input sequences

Retrieved sequences
U/U   U/E
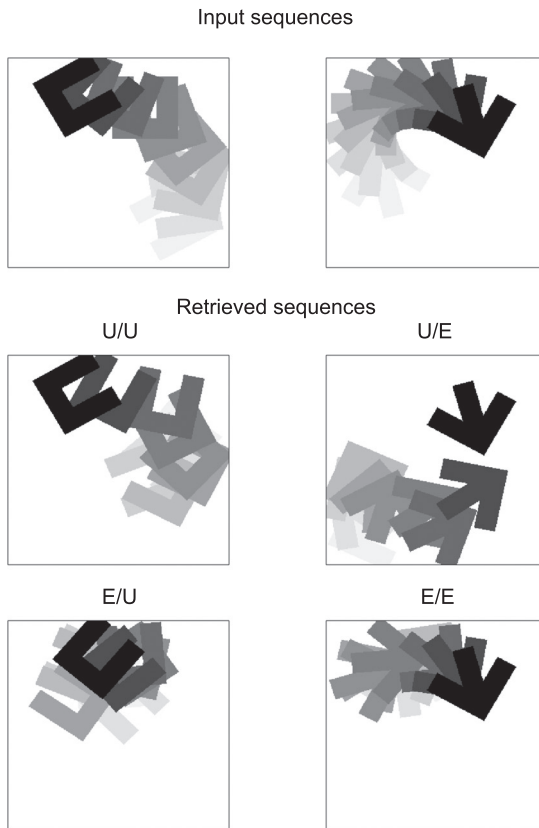

E/U   E/E

Figure 7: Example of stored and retrieved sequences. Each panel contains 10 consecutive elements of one sequence. The shades of gray indicate different elements in the sequence, with darker shades occuring earlier ($\sigma = 0.2$). Top row: Examples of two image sequences that are stored in the sequence storage network. Middle and bottom row: Retrieved sequences using appropriate or inappropriate representations.

Since in the noise-free retrieval case, errors could occur only if two or more stored patterns are identical, our observation implies that no two patterns stored in the network are identical. This conclusion is not surprising since the input stimuli—gray scale, 30×30 pixel images—are real valued and the Lissajous trajectories were designed to be nonrepeating. Since noise-free retrieval is unrealistic, we turned our attention to retrieval with noise. In all four cases, the retrieval error gradually grows as elements of the sequence are retrieved. At higher levels of retrieval noise, the retrieval errors approach the asymptotic error more rapidly.
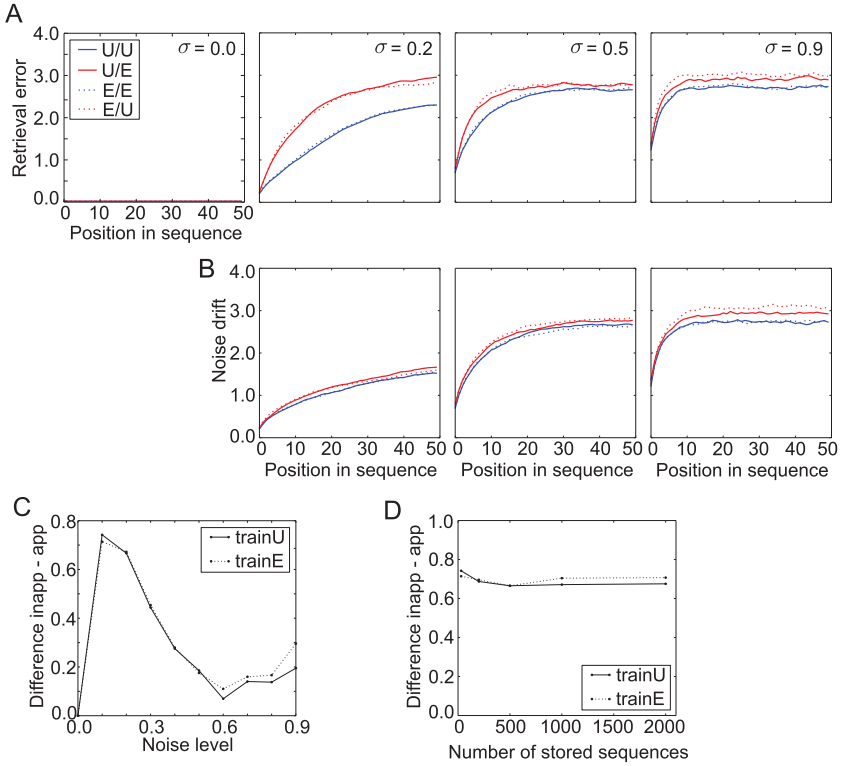
Figure 8: Accuracy of episodic retrieval. (A) Retrieval error for different levels of retrieval noise (given by σ) as a function of the position in the sequence. (B) Accumulation of retrieval noise (noise drift error) for different noise levels corresponding to A. (C) Difference in the retrieval error between appropriate and inappropriate representations as a function of retrieval noise. The retrieval error is assessed at the 30th pattern in the sequence. (D) Summary plot of the difference in the retrieval error between appropriate and inappropriate representations for different memory loads ($\sigma = 0.1$).

Importantly, episodic retrieval with an appropriate semantic representation is more accurate than retrieval with an inappropriate representation for all retrieval noise levels (see Figure 8C). However, the degree of the benefit depends strongly on the retrieval noise. The difference is largest for $\sigma = 0.1$, where retrieval with the appropriate representation is 30% better than with the inappropriate representation and is reduced gradually to lower levels with higher retrieval noise. The nonmonotonic relationship between the retrieval error difference and the retrieval noise level suggests that multiple competing processes influence the retrieval accuracy. Note that in

Figure 8C, we assessed the retrieval error for the 30th sequence element. Since the retrieval error is monotonic in the sequence position, we could have chosen another sequence position for our analysis and obtained similar results. Finally, the difference in retrieval error is very similar for the network trained on the U and the one trained on the E (see Figure 8C), which means that the differences between appropriate and inappropriate representations are not caused by an asymmetry between the two objects.

Another parameter that might affect retrieval performance is the number of stored sequences, since storing more patterns increases the chance of retrieving incorrect patterns. We therefore stored up to 2000 sequences (50 elements each). As expected, the retrieval error increases with the memory load (data not shown), but somewhat surprising, the differences between appropriate and inappropriate representations remain constant (see Figure 8D). We next turn to the causes of the difference in retrieval accuracy.

3.2 Sources of Retrieval Error. Retrieval errors occur when the retrieval process jumps from a correct to an incorrect pattern. Intuitively, one might expect that more frequent jumps to incorrect patterns cause higher retrieval error. Therefore, we first examined the frequency of these jumps. We separately measured the probability of the incorrect jumps that occurs within the same sequence and between different sequences. As expected, the rate of jumps between sequences increases with retrieval noise (see Figure 9A, left). Unexpectedly, the rate was only slightly larger for the inappropriate than for the appropriate representation. The rate of jumps within a sequence was even more surprising. We found that the rate declined for retrieval noise past a certain level, $\sigma > 0.2$, and the rate of jumps within sequences was lower for the inappropriate than for the appropriate representations (see Figure 9A, right). The latter feature contrasts with our previous observation that the retrieval error was larger for the inappropriate representation and seems to indicate that retrieval in the inappropriate case is more robust to noise. The different frequency of incorrect jumps cannot account for the difference in the retrieval error between the two cases. These paradoxical findings warrant further investigation. Our hypothesis is that although the frequency of incorrect jumps is equal or lower when using inappropriate representations, the size of jumps or the subsequent errors, or both, are larger, so that the overall retrieval error is larger on average.

We therefore analyzed different types of errors that may contribute to the retrieval errors. In our model, three types of errors can occur during episodic retrieval. First, the retrieval noise added in each retrieval step can cause incorrect jumps, both within and between sequences (see equation 2.7). These errors accumulate after successive retrievals (noise drift error). Second, our model stores sequences by associating each element with the next element as the retrieval cue. If a jump occurs between sequences, the associated retrieval cue leads to an advancement through the incorrect sequence, which lead retrieval even further away from the original sequence
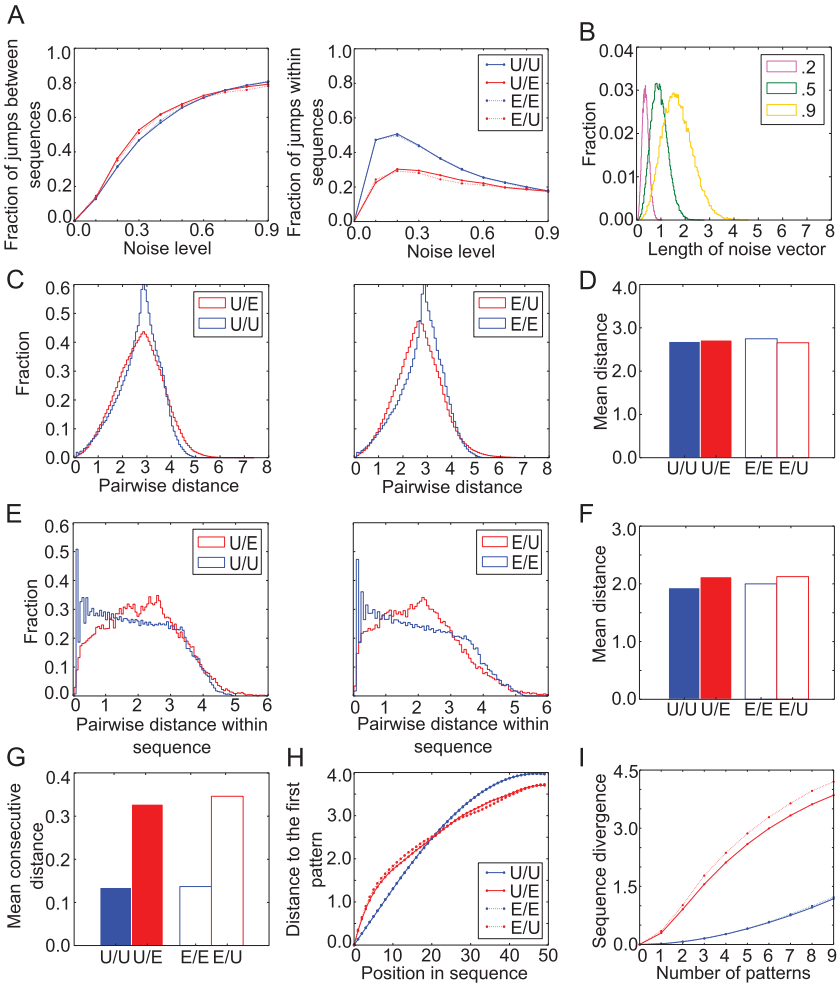
Figure 9: The structure of episodic sequences. (A) Probability of jumping to an incorrect pattern during retrieval. Left: jumps to a different sequence. Right: Jumps within the same sequence. (B) Distribution of retrieval noise added in each retrieval step. Different colors indicate three different noise levels (values indicate σ). (C) Distribution of pairwise distances between any two stored patterns. (D) Mean pairwise distance between any two stored patterns. (E) Distribution of pairwise distances between two patterns within the same sequence. (F) Mean pairwise distance within the same sequence. (G) Mean distance between two consecutive elements in a sequence. (H) Distance of subsequent elements in a sequence from the first pattern. (I) Sequence divergence: the increases in the distance between the subsequent elements of two sequences after two patterns in the respective sequences were the closest patterns to each other.

(sequence divergence error). Third, when two elements are identical, retrieval might proceed in the incorrect sequence. Since the last case does not occur in our simulation (see section 3.1) and since it is highly unlikely that two events repeat precisely in biological systems, we consider only the first two error types in the following.

3.2.1 Noise Drift Error. To isolate the noise drift error from the sequence divergence error, we disable the advancement through the sequence during retrieval. That is, instead of using the associated cue to retrieve the next element in the sequence (see equation 2.7), we used the retrieved pattern itself:

$$s' = \operatorname*{argmin}_{s} \left\| \mathbf{y}'_t + \boldsymbol{\epsilon}_{t+1} - \mathbf{y}_s \right\|. \tag{3.1}$$

With no noise, there would be no drift, and $\mathbf{y}'_t = \mathbf{y}_1$ for all t. We define the noise drift as

$$\eta(t) = \left\| \mathbf{y}'_t - \mathbf{y}_1 \right\|. \tag{3.2}$$

The noise drift is affected by the distribution of patterns stored in the network, which appears to be different between appropriate and inappropriate representations (see Figure 6). Therefore, it might reveal the difference between retrieval with appropriate and inappropriate representations. We found that for large retrieval noise (e.g., $\sigma = 0.9$), the time course of the noise drift is very similar to that of the retrieval error, indicating that the retrieval noise is the dominant contributor to the retrieval error in this case (see Figure 8B). At all noise levels, we also see a difference in noise drift between appropriate and inappropriate representations. However, the difference is small compared to the difference in retrieval error for low retrieval noise ($\sigma = 0.2$). At this noise level, the retrieval error also rises much more steeply than the noise drift.

To better understand these observations, we computed the distribution of distances between two randomly selected patterns. There are clear differences between the distributions for the appropriate and inappropriate representations (see Figure 9C), even though the means are similar (see Figure 9D). We then compared the pairwise distances to the retrieval noise (see Figure 9B). For large retrieval noise ($\sigma = 0.9$), very few noise vectors are longer than 2.5. Since the pairwise distances in this range are more skewed for the inappropriate representation than for the appropriate representation, incorrect jumps occur more often between pairs with relatively larger distance for inappropriate than for appropriate representation, thus accounting for the higher noise drift of the inappropriate representations at high noise levels. By contrast, pairwise distances of patterns within the same sequence (see Figures 9E and 9F) were significantly lower for appropriate

as compared to the inappropriate representations. This explains the unexpected finding that within-sequence jumps are more likely for appropriate representations (see Figure 9A, right). However, since most of these distances are very small, the more frequent within-sequence jumps do not cause larger retrieval errors. Furthermore, the comparison of pairwise distances between patterns in the same sequence and those between all patterns reveals an explanation for the decline of within-sequence jump rates for $\sigma > 0.2$. When the retrieval noise is lower than this level, the noise vectors are shorter than 1 (see Figure 9B). While there are many pair distances from the same sequence within this range (see Figure 9E), they make up only a very small fraction of all pair distances (see Figure 9C). However, since there are also few patterns from other sequences at a short distance, faulty transitions go to patterns from the same as well as other sequences. At larger retrieval noise, however, the noise vector samples from distances that are dominated by patterns from other sequences. Therefore, between-sequence jumps become more likely and the within-sequence jump rate declines (see Figure 9A).

In summary, the distribution of the patterns in pattern space accounts for the properties of the noise drift error and jump probabilities. However, the noise drift errors do not account for the bulk of the difference in the retrieval error between appropriate and inappropriate representations.

3.2.2 Sequence Divergence Error. As a first measure for the structure of the sequences, we considered the consecutive distance: the mean distance between a pattern and the subsequent pattern in the sequence, which is significantly smaller for appropriate representations than for inappropriate representations (see Figure 9G). Figure 10 shows that sequences generated by an inappropriate representation might simply be more jagged, but on average, they follow a similar trajectory as sequences of appropriate representation. This jaggedness or fluctuation itself does not cause a larger retrieval error for inappropriate representations. As we have shown, incorrect jumps driven by noise occur less frequently for inappropriate representations. Higher consecutive distances are consistent with the low fraction of short distance pairs (see Figure 9E), and consequently, low levels of retrieval noise may not suffice to cause retrieval to jump to an incorrect pattern; in other words, retrieval of sequences with higher fluctuation is less sensitive to noise. While considering the interaction between stored sequences during retrieval, the consecutive distance may still increase the retrieval error since once a jump between sequences has occurred, the sequential retrieval process might drive the retrieved pattern farther away from the correct one with larger consecutive distance. We therefore examined the distance of each element from the first element in a sequence. While early in the sequence, the distance to the first element is larger for inappropriate than for appropriate representations, this relationship reverses later in the sequence (see Figure 9H). However, the probability of remaining in the correct
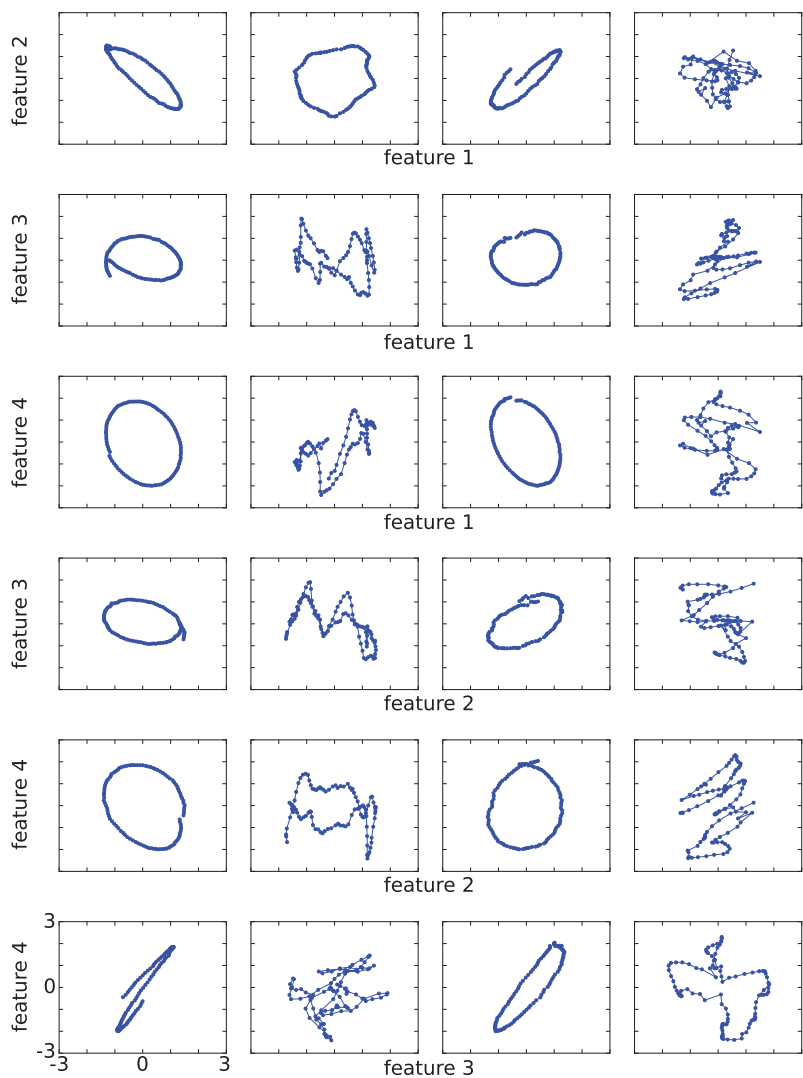
Figure 10: Examples of episodic sequences. The four columns show data for U/U, U/E, E/E, and E/U, respectively. For each case, 100 data points are shown. The 4D pattern vectors are visualized by plotting each dimension against each other.

sequence decreases exponentially with the number of retrieval steps. Since the probability of a between-sequence jump is appreciable even for small amounts of retrieval noise (see Figure 9A, $\sigma > 0.1$), it is highly unlikely that

the retrieval stays in the correct sequence for more than about 10 elements. Therefore, only the distances early in the sequence are relevant for the retrieval performance, and those are larger for the inappropriate representation (see Figure 9H).

To fully account for the larger retrieval error for the inappropriate representations, we need to investigate the relationship between elements in different sequences. Here, we study the sequence divergence, which refers to the tendency of two sequences that are close to each other at some point in time to diverge from each other over time. For a randomly drawn pattern $\mathbf{y}_{p,i}$ (the ith pattern in the pth sequence) from a selected sequence $\mathbf{Y}_p$, we searched for the most similar pattern across all other sequences, say, $\mathbf{y}_{q,j}$ in sequence $\mathbf{Y}_q$. We then followed these two sequences $\{\mathbf{y}_{p,i}, \mathbf{y}_{p,i+1}, \dots, \mathbf{y}_{p,i+n}\}$ and $\{\mathbf{y}_{q,j}, \mathbf{y}_{q,j+1}, \dots, \mathbf{y}_{q,j+n}\}$ and studied the distance between subsequent elements in the two sequences:

$$\xi(k) = \left\| \mathbf{y}_{p,i+k} - \mathbf{y}_{q,j+k} \right\| - \left\| \mathbf{y}_{p,i} - \mathbf{y}_{q,j} \right\|. \tag{3.3}$$

Sequences based on inappropriate representations have higher divergence than those based on appropriate representations (see Figure 9I).

In summary, the higher retrieval error for the inappropriate representation can be interpreted as follows. Although the incorrect jumps occur equally or less often, once a jump between sequences has occurred, retrieval proceeds along the incorrect sequence, which diverges more for the inappropriate representation from the original one, thus leading to larger retrieval errors in comparison with the appropriate representation. In other words, it is the relationship between stored sequences rather than the structure of individual sequences that determines the difference in retrieval performance. This account can be tested by correlating the sequence divergence with the retrieval error. Since the four data points we have obtained so far are insufficient for such an analysis and to show that our results generalize to objects other than the U and E, we next turn to analyzing other objects in the same framework.

3.3 Generalization to Other Objects and Classes of Objects. To show that our findings above hold in general, we study episodic retrieval with other objects (T versus L and H versus X); (see Figures 11A and 11B), and since categorization is an important feature in semantic memory, with two arbitrary categories of objects (T and L versus U and E); (see Figure 11C). The simulations and analyses were identical to the ones reported above except where otherwise noted explicitly. In particular, since the letters H and X have a rotational symmetry that the other objects lack, we set the rotating speed of these two letters to half the value that we used for the other objects. For training with categories, the two objects appear one at a time and alternate with a transition probability of 0.05 in each step. There are no blank
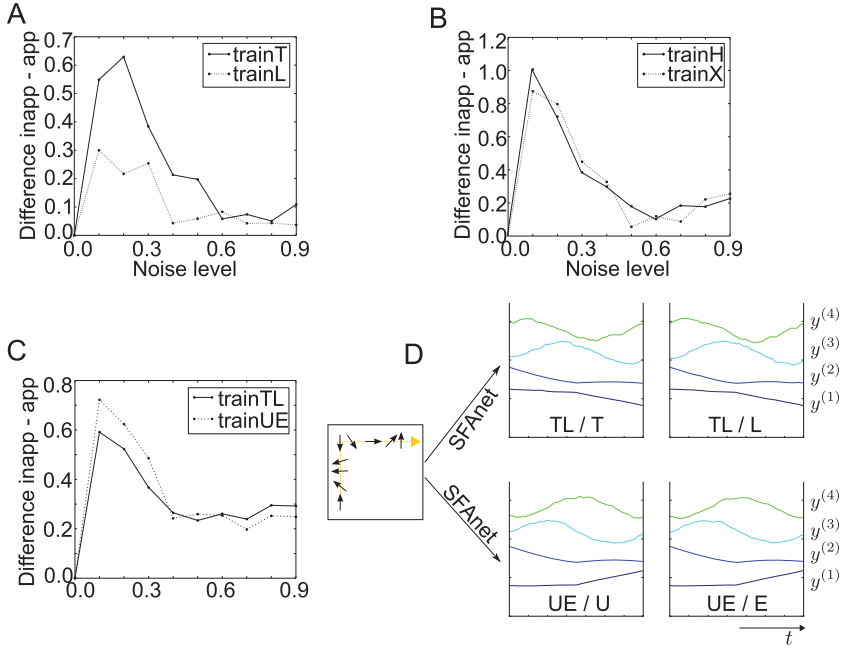
Figure 11: Difference in retrieval error between appropriate and inappropriate representations generalizes to other objects and groups of objects. The retrieval error for appropriate representations is lower than for inappropriate representations when using (A) objects T and L, (B) objects H and X, and (C) two categories of objects: {U, E} and {T, L}. (D) In the last case, SFA learns common features for the two objects within the same category. Plotting convention as in Figure 4.

images at the time when the two objects alternate. This ensures that SFA extracts features that are independent of the objects (see Figure 11D). In the testing phase, we examined the retrieval performance on each object separately and averaged across both objects to report the results for the category. All new analyses confirm that the episodic retrieval is more accurate with appropriate representations than with inappropriate representations (see Figures 11A to 11C). Moreover, we verified that our results are not artifacts of the Lissajous trajectories that we use to generate sequences by obtaining very similar results with random walk trajectories and rotation (data not shown).

3.3.1 *Differences in Sequence Divergence Account for Differences in Retrieval Quality.* With these additional data points, we can plot the retrieval error against the sequence divergence (see Figure 12). Note that we did not average the results across the two objects for the group simulation. We found
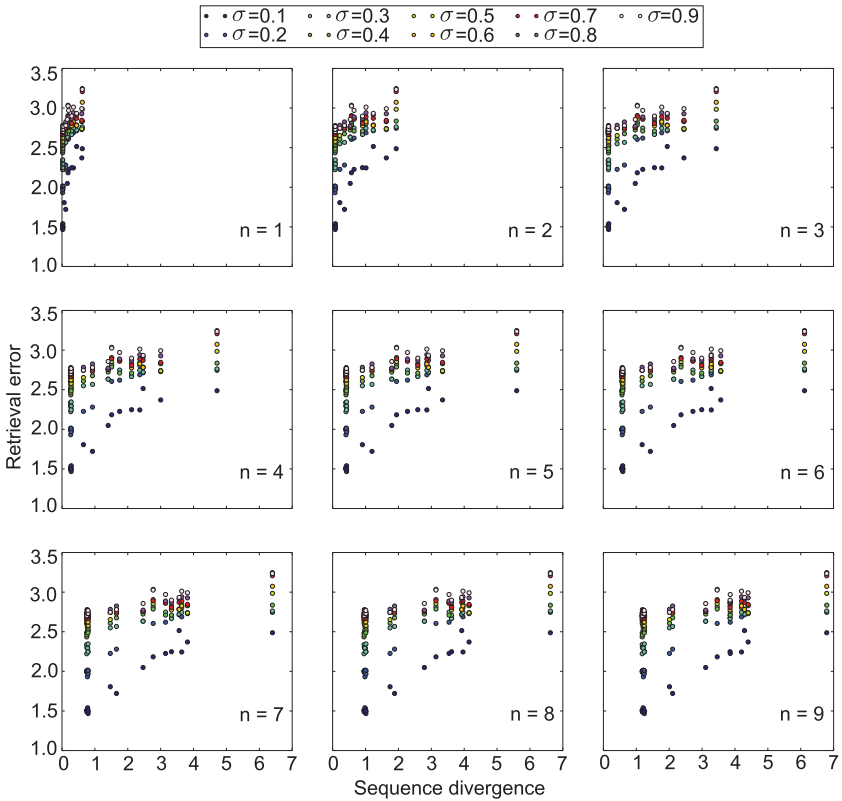
Figure 12: The sequence divergence monotonically increases with the retrieval error. Each panel summarizes the relationship between the two variables after a particular number of retrieval steps (n). Different retrieval noise levels are indicated by different colors (σ).

that the relationship increases sharply at low noise levels and saturates at higher noise levels. Most important for us, the relationship is monotonic, showing that differences in sequence divergence indeed account for the differences in retrieval quality.

4 Experimental Study

Although this study is primarily a modeling study, we explore whether the model's prediction has any grounding in reality. Since we were not aware of any experimental study in the literature that would be compatible with our modeling, we decided to conduct a pilot experiment to test the prediction of our computational model in principle. That is, we studied whether the

facilitating effect of the semantic representation on episodic memory can be observed experimentally. Since human subjects have had a tremendous amount of visual experience with objects, we reasoned that an exposure equivalent to that in the model would not be sufficient to alter their semantic representations. Conversely, for technical reasons, we could not train our model on data as rich as experienced by humans. We therefore designed an experiment in which subjects had the opportunity to improve their semantic representations of certain objects, but not others, in a different fashion and then tested their memory for episodes involving both types of objects.

4.1 Methods. The study consisted of two phases on two consecutive days: a semantic training on day 1 and an episodic memory test on day 2.

4.1.1 Participants. Nineteen (10 female, 9 male) healthy subjects with a mean age of 26.3 years (SD = 6.0 years; range: 19 to 42 years) participated in the experiment. They were recruited among the psychology students at the Ruhr University Bochum and received course credit for their participation. All participants gave written informed consent, and the study was conducted in accordance with the principles of the Declaration of Helsinki. The study was approved by the ethics committee of the Faculty of Psychology, Ruhr University Bochum.

4.1.2 Stimuli and Semantic Training. Eight different shapes, which did not resemble real objects or living things, served as stimuli (see Figure 13A). The total group of stimuli was divided into two sets, A and B, consisting of four stimuli each. During the training phase on day 1, subjects learned information about one set of objects (the trained set), whereas the other set was not part of the training (untrained set). The assignment of the object sets (A and B) to the training conditions was varied among subjects.

Since adult humans have already acquired the capability to extract simple object features, we cannot use the same training procedure as we did for the model. The main requirement is that subjects have a more appropriate representation for the trained objects than for the untrained objects. How this difference came about is secondary for the purpose of the study. We therefore devised a training procedure that we reasoned to improve a subject's semantic representation of the trained objects. Subjects looked at the training objects and touched exemplars of the objects that were cut out of black board. They could hold each object in their hands for 10 seconds and were asked to try to remember the shape of the objects. Subjects were then given so-called ID cards for each object, which contained a picture of the object's shape together with four types of information that the subjects were asked to learn: the name (e.g., Christoph), the favorite color (e.g., red), the favorite food (e.g., spinach), and one ability (e.g., digging). Subjects were handed the four ID cards for the trained objects one at a time for a duration of 150 seconds (total duration 600 seconds). The instruction was to
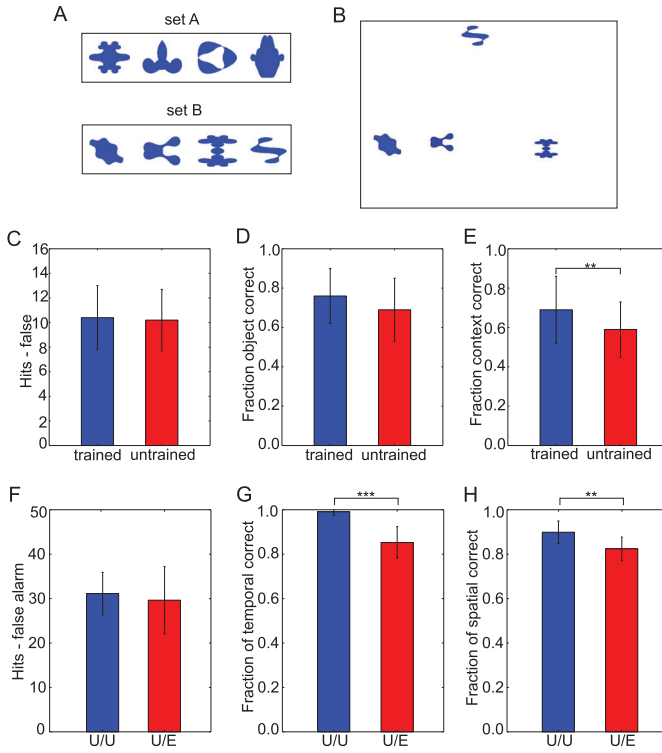
Figure 13: Stimuli, results and model simulation of an episodic memory experiment using familiar and novel objects. (A) Two sets of objects were used as stimuli in the episodic memory test. A given subject received semantic pre-training on one set of objects only (counterbalanced across subjects), but subjects' episodic memory was tested on both sets of objects. (B) A screenshot from one of the video clips used to test episodic memory. In the video, objects are moving around and can appear, disappear, and collide with other objects. (C) Subjects' performance score for remembering the occurrence of events involving the trained and untrained objects ($p = .824$, t-test). (D) Fraction with which subjects correctly remembered objects out of all remembered events ($p = .172$, t-test). (E) Fraction of correctly remembered context information out of all remembered events. Subjects show significantly better performance on trained than on untrained stimulus sets ($p < 0.01$, t-test). (F) Model performance on the occurrence of events (50 trials in total, average across 30 repetitions, comparing the appropriate representation (U/U) to the inappropriate representation (U/E) ($p = .437$, t-test). (G) Fraction of correctly retrieved temporal information out of all retrieved events ($p < 0.001$, t-test). (H) Fraction of correctly retrieved spatial information out of all retrieved events. For both temporal and spatial information, our model shows significantly better performance with an appropriate semantic representation (U/U) than with an inappropriate one (U/E) ($p < 0.01$, t-test). Error bars indicate standard deviations.

Table 1: Potential Events in the Videos during the Test Phase.

	Events
One object	Appears
	Moves in a zigzag manner
	Touches a border of the screen
	Moves parallel to a side border of the screen
	Stops moving
	Disappears
Two objects	Move in parallel
	Touch each other
Three objects	Are aligned along a line; the fourth one is not

memorize the information for a subsequent test, which followed after the last ID card and a short break of about 5 minutes. This semantic test contained 16 questions, one for each piece of information for each object (e.g., "Which of the objects can fly?"; "What is the favorite food of this object?"). For each question, subjects could choose among four different alternatives, exactly one of which was correct. Importantly, participants had to reach a performance criterion of 12 correct answers in order to be admitted to the test phase on day 2. If they did not reach this criterion, the procedure of the learning phase was repeated and the test was conducted a second time.

4.1.3 Episodic Memory Test. On day 2, subjects were shown eight videos of 30 s duration each. The videos showed movements of either the four trained or the four untrained objects (four videos each; a screenshot is shown in Figure 13B). Based on the objects' movements, different events could occur in each video (see Table 1 for a list of potential events).

Before the first test video was shown, subjects were informed about the types of events that could occur and instructed to memorize the events as accurately as possible. To familiarize them with the task, participants were first shown a training video that contained well-known shapes (e.g., circle) and some of the potential events listed in Table 1.

Following each of the eight test videos, subjects received a questionnaire in which their episodic memory for the events in the preceding video was tested. In the questionnaires, eight events were listed, four of which occurred in the preceding video and four of which did not ("old" versus "new"). Subjects were first asked, for each event, whether it had occurred in the video. If subjects answered positively, they were asked which one of the four objects was involved in this event. In case of an event involving multiple objects, subjects had to choose among four different combinations of objects. The last question for a given event referred to the context—either spatial (four of the mentioned events per questionnaire) or temporal (the remaining four). For the questions concerning the spatial context, subjects

were asked to indicate the quadrant of the screen in which quadrant the event occurred by marking the quadrant of a rectangle on the questionnaire. Similarly, questions about the temporal context of an event asked in which quarter interval the event had occurred by marking the respective section on a time bar. Videos with trained and untrained objects alternated, with the type of stimuli in the first video being counterbalanced between subjects. To quantify each subject's performance in the test phase, the sums of hits (i.e., correctly recognized events), false alarms (wrong answers to events that did not occur), correctly identified objects for an event, and correct context memories were calculated separately for trained and untrained objects. After the completion of all questionnaires, the semantic test from day 1 was repeated in order to test whether the newly acquired knowledge was still available on the test day.

4.1.4 Data Analysis. Three dependent variables were considered in the analysis of the subjects' performance in the episodic memory test, separately for videos with trained and untrained objects. As a general performance measure, the sum of false alarms in the evaluation of events (old versus new) was subtracted from the sum of hits. Then the percentage of events, for which the associated objects were remembered correctly, was calculated relative to the number of correctly remembered events. Finally, context knowledge was assessed as the percentage of correct responses on context questions relative to the number of correctly remembered events. For all three variables, paired *t*-tests were performed, comparing the measures for trained and untrained objects. *P* values below 0.05 were considered significant.

4.2 Experimental Confirmation of Our Theoretical Prediction. On average, the subjects' performance score for remembering the events in the videos (hits-false alarms, maximum score 16) amounted to 10.4 (SD = 2.6) for trained objects and 10.2 (SD = 2.5) for untrained objects. The difference between conditions was not significant ($p = .824$). For the remembered events, subjects remembered trained objects as well as untrained objects (mean trained: 76%, SD = 14%; mean untrained: 69%, SD = 16%; $p = .172$). However, a significant difference emerged when memory for context information was tested. With a score of 69% (SD = 17%), context memory was significantly better for events involving trained objects than for events with untrained objects (mean 59%, SD = 14%; t(18) = 2.949, $p = .009$) (see Figures 13C to 13E). The semantic test, administered after the completion of the last episodic memory questionnaire on day 2, yielded an average score of 14.2 (SD = 2.5). Only three participants scored lower than 12, which was the cutoff criterion for the semantic test after the training session on day 1. Exclusion of these three subjects from the analysis yielded the same pattern of results. Subjects had indeed retained the information learned during the

semantic training, suggesting that the semantic representations of the trained objects might be more appropriate than of the untrained objects.

The behavioral results suggest that some key aspects (spatial and temporal) of episodic memory are improved by having a more appropriate semantic representation of the objects involved, even if not all aspects of episodic memory benefited from semantic training. We return to this point in the section 5.

4.3 Modeling Our Experimental Results. The retrieval error is a good diagnostic measure for the performance of a computational model, but it cannot be compared directly to experimental results. We therefore augmented our modeling with postretrieval mechanisms to be able to compare the model directly to our experimental results. From the events that occured in the videos shown to human subjects (see Table 1), we modeled the detection of the appearance and disappearance events since these could be implemented with the fewest additional assumptions. In each sequence, the moving object can either appear or disappear once at a given time. When present, the object moves along a random walk trajectory and rotates simultaneously. When the object is absent, the inputs consist of blank images. The task of the model, not unlike that of the human subjects, is to detect the event in the output sequences of SFA and, if an event is detected, determine the context information of the event, including temporal and spatial information. Since the SFA network generates a constant output when the image is blank and time-varying representations for moving objects, we chose to detect the event by examining the time derivative of the stored sequences. If the derivative is above a threshold, the model reports an event. This task is not trivial when retrieval noise $\epsilon_t^{(k)} \sim N(0, \sigma^2)$ is added during recall, as we do in our other simulations.

The number of trials with correct event detection is the number of hits. General performance is measured by subtracting the sum of false alarms (the number of trials the model detects an event, even though no event had occurred) from the number of hits. To determine the time of the event, we divided the sequence into four equal time windows. The time window, in which the event is detected, is reported by the model. To extract the spatial information about the event, we note that the SFA network extracts features corresponding to the spatial information (x-, y-coordinates) of the moving object. We divided the 2D input space into four quadrants and asked which corner the SFA output was closest to when the event occurred. We then measured the percentage of the temporal and spatial information that the model reports correctly.

In this way, our model is able to partially replicate the experimental results. While there is no significant difference in the performance of detecting the occurrence of an event between appropriate and inappropriate conditions (see Figure 13F), there were significant differences in retrieving

the time (see Figure 13G) and spatial location (see Figure 13H) of the event. These results derive from the fact that SFA generates more jagged sequences with inappropriate representations and therefore information of the event would be detected less precisely by measuring the time derivative. Although our model shows overall better performance on remembering context information than human subjects do, our results are consistent with our experimental findings. Taken together, these results suggest that appropriate semantic representations improve episodic memory but that improvement is not always apparent in every behavioral variable.

5 Discussion

We have presented and studied a model of the interaction between the episodic and semantic system. The main assumptions of our model are that (1) all incoming information must be processed first by the semantic system before it is encoded and stored into episodic memory; (2) semantic memory is represented as static neural patterns, which represent slowly varying features that are hidden in the raw sensory input; (3) episodic memory is represented as temporal sequences of patterns, which are semantic representations of the sensory input; and (4) episodic memory cannot be retrieved directly but rather relies on the semantic system. We found in our model that episodic retrieval is more accurate if the stored sequences were encoded with an appropriate semantic representation. We emphasize that we did not attempt to model the neural mechanisms of semantic representation and episodic memory storage per se but, rather, the effective output of the computations performed by these systems. Consistent with our model's prediction, human subjects remember contextual information better if the episodes involved trained objects, for which they arguably had developed a better semantic representation. Overall, our results suggest a specific computational function for the interaction between the episodic and semantic systems, which helps clarify their disputed relationship.

5.1 Semantic Representation as an Interface to and for Episodic Memory. Serial encoding is central in our model, but not universally accepted. For instance, Simons, Graham, and Hodges (2002) suggested that external information can also enter the episodic system directly through the perceptual system. In other words, varying the contribution from either the semantic system or perceptual system individually has little effect on episodic memory (Graham et al., 2000; Simons & Graham, 2000; Simons, Graham, Galton, Patterson, & Hodges, 2001; Mayes & Roberts, 2001). While Tulving (2001) criticized the methodology of these studies, our model suggests an alternative interpretation that reconciles those results with serial encoding. In our model, the semantic system can represent inputs that it was never trained on using an inappropriate semantic representation. The accuracy of

episodic memory in this case is even quite reasonable, just not as high as for memories stored with an appropriate representation.

In our study, the semantic system was trained only once and then remained static throughout testing. In principle, our model can accommodate changes in the semantic system after episodic memories have been stored. Such changes will lead to distortions of the episodic memories. Other models have found that episodic memory is fragile and might be easily eroded in the face of neocortical plasticity (Káli & Dayan, 2004). If the semantic system changes severely, it might become impossible to recall the stored episodes, as happens perhaps in childhood amnesia. Children can remember events from before the age of 3 or 4 years, but these memories fade away as children get older and adults show near total forgetting of memories from their early childhood (Fivush & Schwarzmueller, 1998; Rubin, 2000). Future work is needed to study the distortion or forgetting of episodic memory due to changes in the semantic system. Another perspective for future work is that semantic information can be acquired from learned episodes as well, which means that the semantic representation layer could be trained by episodic memory.

5.2 Chaining Model and Episodic System. The chaining model was a popular approach to model serial-order memory (Murdock, 1993, 1995; Lewandowsky & Murdock, 1989). However, a number of studies have demonstrated over the years that chaining does not account for some key features of serial order in short-term memory (Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Henson, 1996a, 1998; Page & Norris, 1998). As a result, contemporary computational studies turn to other principles (e.g., Burgess & Hitch, 1999; Farrell, 2012; Henson, 1998; Page & Norris, 2009). One essential distinction of our study from these other ones is that we model the episodic memory system, not sequence learning in general, and, more specifically, the role of CA3 in episodic memory. Our algorithmic model stores the association between consecutive memory patterns in a sequence, though this association in our model is predetermined rather than built during learning. By contrast, most cognitive studies of serial-order recall have been performed in the framework of working memory (see Hurlstone, Hitch, & Baddeley, 2014, for a review). For instance, subjects are typically asked to recall the list immediately after its last item disappeared (e.g., Henson, 1996b), which can hardly be considered episodic retrieval in our view. The relationship between episodic memory and working memory is beyond the scope of our current study.

5.3 Comparison to Other Computational Models. While numerous experimental studies have investigated the relation between different memory systems, very few computational studies have done so. Howard and colleagues (Howard & Kahana, 2002; Howard, Shankar, & Jagadisan, 2011) suggest that it is possible to build a model of semantic memory acquisition

in the same framework, involving drifting temporal context, as their model of episodic recall. Káli and Dayan (2004) have suggested that episodic memory and semantic memory share the same representational format (i.e., static patterns) but are distinguished by different amounts of overlap between different memory patterns. However, their focus was on studying the effect of replay on memory in the cortico-hippocampal loop, and so they did not model the storage of episodic memories in the hippocampus. Hintzman's Minerva 2 model proposed that two distinct memory types can be produced by a model that only stores episodes (Hintzman, 1984, 1986). The model is a multiple trace model that stores all repetitions of an item as different episodic traces. Correspondingly, the information retrieved reflects the summed content of all activated traces in parallel. Abstract concepts, or semantic information, can be derived from the pool of episodic traces as an artifact of averaging at retrieval ("prototype effect"). Finally, Battaglia and Pennartz (2011) modeled the formation of semantic memory from relational episodic information during memory consolidation. That study examined the opposite flow of information as compared to our computational model, from the episodic to the semantic system.

Our conceptual and algorithmic model of episodic memory is fully consistent with and complementary to a theoretical framework for the function of the hippocampus in episodic memories that was proposed by Cheng (2013). That framework, called CRISP, proposes that the recurrent connections in region CA3 generate intrinsic neural sequences, which are then associated with sequences of external inputs to store episodic memories. While the complete CRISP framework has not been implemented in a computational model yet, some aspects such as the generation of intrinsic sequence in CA3 have been (Azizi, Wiskott, & Cheng, 2013; Hopfield, 2009). A central prediction of CRISP, that pattern completion in the hippocampus might not depend on CA3, has been confirmed in a computational study (Neher, Cheng, & Wiskott, 2015). In our study, we use a simplified algorithmic model to store and retrieve episodic sequences. Now that we have a better grasp of how this sequence storage interacts with the semantic system, the next step will be to implement sequence storage in a neural network.

5.4 Experimental Test of Our Theoretical Predictions. While the details of our modeling study and our experiment are not well matched, they follow parallel designs in principle. Subjects/networks first received semantic training on a certain set of objects; then episodic memory was tested on the training stimuli as well as unknown ones. This way, we can use a within-subject design to compare episodic retrieval performance when the semantic representation is appropriate to when it is inappropriate. We found better episodic memory for contextual information for the trained than the untrained objects, but comparable performance on the occurrence of events in both human subjects and our computational model. Two additional aspects are worth keeping in mind when interpreting the differing

effects of semantic representations on different measures of episodic memory. First, our adult subjects have had many exposures to all sorts of objects during their lifetime, for which they have developed semantic representations. The training in our experiment is expected to improve the semantic representation of the trained objects by only a small amount relative to the representation of the untrained objects. Hence, only a small effect would be expected on episodic memory. Second, the semantic training (exposure to the object and uncorrelated information) and testing was unrelated to the episodic memory task (remembering video clips). Any improvement on episodic memory is an effect of generalization and therefore expected to be small. These small improvements might be observable in only some measures of memory performance.

Other studies have shown that a familiarization process similar to the one in our experiment facilitates item recognition memory, free recall, and associative memory (Ratcliff, Clark, & Shiffrin, 1990; Murnane & Shiffrin, 1991; Kilb & Naveh-Benjamin, 2011). These effects can be introduced through either longer exposure time or repetition of items during studying phase. Although the underlying mechanism is still not clear, McClelland and Chappell (1998) suggested that familiarization with an item induces more accurate knowledge of its characteristics, which allows it to be easily differentiated from the others. This interpretation might be compatible with our view that the semantic training in our study allowed subjects to build a more appropriate semantic representation of the trained objects.

Nevertheless, future studies are needed to bridge the differences between our current modeling and experimental setups. For one, the training paradigms were quite different. Although the sole purpose of the training was to improve the semantic representations and not to learn particular content, it would nonetheless be preferable to have more similar training procedures. For two, multiple objects appear simultaneously in the experiment and interacted in the videos, while only one object was visible at a time in our computational studies. Simpler sequences than those used in the experiment would be too simple for human subjects. More complex image sequence could not be analyzed in our current model. However, in principle, SFA can be used to extract object identity from a group of training objects displayed alternatively (Franzius et al., 2011) or simultaneously (Legenstein, Wilbert, & Wiskott, 2010). Hence, an interesting future study would be to test whether our results hold for features including object identity.

6 Conclusion

We have proposed an abstract computational model of the interdependence between the semantic and episodic system. In this study, we have presented evidence from both computational and experimental studies suggesting

that semantic representations play an important role for the encoding and retrieval of episodic memory, in particular, the spatial and temporal aspects. Our model is able to cope with high-dimensional visual input and store and retrieve temporal sequences. Such properties make it different from other models. However, many aspects of our model need to be investigated further by both experimental and computational studies, ideally using parallel designs.

Acknowledgments

References

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, *17*(1), 1–12.

Azizi, A. H., Wiskott, L., & Cheng, S. (2013). A computational model for preplay in the hippocampus. *Frontiers in Computational Neuroscience*, *7*, 161.

Baddeley, A. (1984). Neuropsychological evidence and the semantic/episodic distinction. *Behavioral and Brain Sciences*, *7*, 238–239.

Bartlett, F. C., & Kintsch, W. (1932). *Remembering*. Cambridge: Cambridge University Press.

Battaglia, F. P., & Pennartz, C. M. A. (2011). The construction of semantic memory: Grammar-based representations learned from relational episodic information. *Frontiers in Computational Neuroscience*, *5*, 36.

Bayley, P. J., Hopkins, R. O., & Squire, L. R. (2006). The fate of old memories after medial temporal lobe damage. *Journal of Neuroscience*, *26*(51), 13311–13317.

Berkes, P., & Wiskott, L. (2005). Slow feature analysis yields a rich repertoire of complex cell properties. *Journal of Vision*, *5*(6), 572–602.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.

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

Buhry, L., Azizi, A. H., & Cheng, S. (2011). Reactivation, replay, and preplay: How it might all fit together. *Neural Plasticity*, *2011*, 1–11.

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.

Burianova, H., & Grady, C. L. (2007). Common and unique neural activations in autobiographical, episodic, and semantic retrieval. *Journal of Cognitive Neuroscience*, *19*(9), 1520–1534.

Burianova, H., McIntosh, A. R., & Grady, C. L. (2010). A common functional brain network for autobiographical, episodic, and semantic memory retrieval. *NeuroImage*, *49*(1), 865–874.

Cabeza, R., & Nyberg, L. (2000). Neural bases of learning and memory: Functional neuroimaging evidence. *Current Opinion in Neurology*, *13*(4), 415–421.

Chan, D., Fox, N. C., Scahill, R. I., Crum, W. R., Whitwell, J. L., Leschziner, G., . . . Rossor, M. N. (2001). Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Annals of Neurology*, *49*(4), 433–442.

Cheng, S. (2013). The CRISP theory of hippocampal function in episodic memory. *Frontiers in Neural Circuits*, *7*, 88.

Cheng, S., & Werning, M. (2016). What is episodic memory if it is a natural kind? *Synthese*, *193*(5), 1345–1385.

Cheng, S., Werning, M., & Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. *Neuroscience and Biobehavioral Reviews*, *60*, 82–89.

Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003). Opinion: Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, *4*(8), 685–691.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, *47*(11), 2305–2313.

Craik, F., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

Craik, F., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Duval, C., Desgranges, B., de La Sayette, V., Belliard, S., Eustache, F., & Piolino, P. (2012). What happens to personal identity when semantic knowledge degrades? A study of the self and autobiographical memory in semantic dementia. *Neuropsychologia*, *50*(2), 254–265.

Düzel, E., Habib, R., Guderian, S., & Heinze, H. J. (2004). Four types of novelty-familiarity responses in associative recognition memory of humans. *European Journal of Neuroscience*, *19*(5), 1408–1416.

Ebbinghaus, H., Ruger, H. A., & Bussenius, C. E. (1913). *Memory: A contribution to experimental psychology.* New York. Dover. (Trans. H. A. Ruger & C. E. Bussenius. Original work published 1885).

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*, 223–271.

Fivush, R., & Schwarzmueller, A. (1998). Children remember childhood: Implications for childhood amnesia. *Applied Cognitive Psychology*, *12*(5), 455–473.

Franzius, M., Sprekeler, H., & Wiskott, L. (2007). Slowness and sparseness lead to place, head-direction, and spatial-view cells. *PLoS Computational Biology*, *3*(8), e166.

Franzius, M., Wilbert, N., & Wiskott, L. (2011). Invariant object recognition and pose estimation with slow feature analysis. *Neural Computation*, *23*(9), 2289–2323.

Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences*, *95*(3), 906–913.

Gilboa, A. (2004). Autobiographical and episodic memory—one and the same? Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia*, *42*(10), 1336–1349.

Graham, K. S., & Hodges, J. R. (1997). Differentiating the roles of the hippocampal complex and the neocortex in long-term memory storage: Evidence from the study of semantic dementia and Alzheimer's disease. *Neuropsychology*, *11*(1), 77–89.

Graham, K. S., Kropelnicki, A., Goldman, W. P., & Hodges, J. R. (2003). Two further investigations of autobiographical memory in semantic dementia. *Cortex*, *39*(4–5), 729–750.

Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia*, *38*(3), 313–324.

Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, *16*(5), 748–753.

Hemmer, P., & Steyvers, M. (2009). Integrating episodic memories and prior knowledge at multiple levels of abstraction. *Psychonomic Bulletin and Review*, *16*(1), 80–87.

Henson, R. N. A. (1996a). *Short-term memory for serial order*. PhD diss., Cambridge University.

Henson, R. N. A. (1996b). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology Section A*, *49*(1), 80–115.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, *36*(2), 73–137.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, and Computers*, *16*(2), 96–101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428.

Hopfield, J. J. (2009). Neurodynamics of mental exploration. *Proceedings of the National Academy of Sciences*, *107*(4), 1648–1653.

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*(1), 85–98.

Howard, M. W., Shankar, K. H., & Jagadisan, U. K. K. (2011). Constructing semantic representations from a gradually-changing representation of temporal context. *Topics in Cognitive Science*, *3*(1), 48–73.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339–73.

Huth, A., Nishimoto, S., Vu, A., & Gallant, J. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210–1224.

Irish, M., Addis, D. R., Hodges, J. R., & Piguet, O. (2012). Considering the role of semantic memory in episodic future thinking: Evidence from semantic dementia. *Brain*, *135*, 2178–2191.

Jensen, O., & Lisman, J. E. (1996). Hippocampal CA3 region predicts memory sequences: Accounting for the phase precession of place cells. *Learning and Memory*, *3*(2–3), 279–287.

Jones, D. M., Beaman, P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 209–238). New York: Psychology Press.

Káli, S., & Dayan, P. (2004). Off-line replay maintains declarative memories in a model of hippocampal-neocortical interactions. *Nature Neuroscience*, *7*(3), 286–294.

Kan, I. P., Alexander, M. P., & Verfaellie, M. (2009). Contribution of prior semantic knowledge to new episodic learning in amnesia. *Journal of Cognitive Neuroscience*, *21*(5), 938–944.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1998). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Mijake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.

Kilb, A., & Naveh-Benjamin, M. (2011). The effects of pure pair repetition on younger and older adults' associative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 706–719.

Kinsbourne, M., Rufo, D. T., Gamzu, E., Palmer, R. L., & Berliner, A. K. (1991). Neuropsychological deficits in adults with dyslexia. *Developmental Medicine and Child Neurology*, *33*(9), 763–775.

Kitchener, E. G., Hodges, J. R., & McCarthy, R. (1998). Acquisition of post-morbid vocabulary and semantic facts in the absence of episodic memory. *Brain*, *121*(7), 1313–1327.

Klein, S. B. (2013a). Making the case that episodic recollection is attributable to operations occurring at retrieval rather than to content stored in a dedicated subsystem of long-term memory. *Frontiers in Behavioral Neuroscience*, *7*, 3.

Klein, S. B. (2013b). The complex act of projecting oneself into the future. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(1), 63–79.

Kopelman, M. D., Stanhope, N., & Kingsley, D. (1999). Retrograde amnesia in patients with diencephalic, temporal lobe or frontal lesions. *Neuropsychologia*, *37*(8), 939–958.

Lashley, K. S. (1951). *The problem of serial order in behavior.* New York: Wiley.

Legenstein, R., Wilbert, N., & Wiskott, L. (2010). Reinforcement learning on slow features of high-dimensional input streams. *PLoS Computational Biology*, *6*(8), e1000894.

Levy, W. B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus*, *6*(6), 579–590.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, *96*, 25–57.

Lisman, J. E. (1999). Relating hippocampal circuitry to function: Recall of memory sequences by reciprocal dentate-CA3 interactions. *Neuron*, *22*(2), 233–242.

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review*, *109*(2), 376–400.

Maguire, E. A. (2001). Neuroimaging studies of autobiographical event memory. *Philosophical Transactions of the Royal Society B*, *356*(1413), 1441–1451.

Maguire, E. A., Kumaran, D., Hassabis, D., & Kopelman, M. D. (2010). Autobiographical memory in semantic dementia: A longitudinal fMRI study. *Neuropsychologia*, *48*(1), 123–136.

Mangels, J. A. (1997). Strategic processing and memory for temporal order in patients with frontal lobe lesions. *Neuropsychology*, *11*(2), 207–221.

Manns, J. R., Hopkins, R. O., & Squire, L. R. (2003). Semantic memory and the human hippocampus. *Neuron*, *38*(1), 127–133.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, *11*(2), 194–201.

Mayes, A. R., & Roberts, N. (2001). Theories of episodic memory. *Philosophical Transactions of the Royal Society B*, 356(1413), 1395–1408.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, *105*(4), 724–760.

Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., . . . Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*, *133*(11), 3256–3268.

Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review*, *100*(2), 183–203.

Murdock, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory and Cognition*, *23*(5), 631–645.

Murnane, K., & Shiffrin, R. M. (1991). Word repetitions in sentence recognition. *Memory and Cognition*, *19*(2), 119–130.

Neher, T., Cheng, S., & Wiskott, L. (2015). Memory storage fidelity in the hippocampal circuit: The role of subregions and input statistics. *PLoS Computational Biology*, *11*(5), e1004250.

Nyberg, L., Forkstam, C., Petersson, K. M., Cabeza, R., & Ingvar, M. (2002). Brain imaging of human memory systems: Between-systems similarities and within-system differences. *Cognitive Brain Research*, *13*(2), 281–292.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*(4), 761–781.

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society B*, *364*(1536), 3737–3753.

Rajah, M. N., & McIntosh, A. R. (2005). Overlap in the functional neural systems involved in semantic and episodic memory retrieval. *Journal of Cognitive Neuroscience*, *17*(3), 470–482.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 163–178.

Rosenbaum, R. S., Moscovitch, M., Foster, J. K., Schnyer, D. M., Gao, F. Q., Kovacevic, N., . . . Levine, B. (2008). Patterns of autobiographical memory loss in medial-temporal lobe amnesic patients. *Journal of Cognitive Neuroscience*, *20*(8), 1490–1506.

Rubin, D. C. (2000). The distribution of early childhood memories. *Memory*, *8*(4), 265–269.

Ryan, L., Cox, C., Hayes, S. M., & Nadel, L. (2008). Hippocampal activation during episodic and semantic memory retrieval: Comparing category production and category cued recall. *Neuropsychologia*, *46*(8), 2109–2121.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REMretrieving effectively from memory. *Psychonomic Bulletin and Review*, *4*(2), 145–166.

Simons, J. S., & Graham, K. S. (2000). New learning in semantic dementia: Implications for cognitive and neuroanatomical models of long-term memory. *Revue de Neuropsychologie*, *10*(1), 199–215.

Simons, J. S., Graham, K. S., Galton, C. J., Patterson, K., & Hodges, J. R. (2001). Semantic knowledge and episodic memory for faces in semantic dementia. *Neuropsychology*, *15*(1), 101–114.

Simons, J. S., Graham, K. S., & Hodges, J. R. (2002). Perceptual and semantic contributions to episodic memory: Evidence from semantic dementia and Alzheimer's disease. *Journal of Memory and Language*, *47*(2), 197–213.

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, *21*(3), 489–510.

Stuss, D. T., Craik, F., Sayer, L., Franchi, D., & Alexander, M. P. (1996). Comparison of older people and patients with frontal lesions: Evidence from word list learning. *Psychology and Aging*, *11*(3), 387–395.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, *123*(2), 133–167.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*, *44*(12), 2189–2208.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*(4), 385–398.

Tulving, E. (1995). Organization of memory: Quo vadis. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 839–847). Cambridge, MA: MIT Press.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society B*, *356*(1413), 1505–1515.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, *8*(3), 198–204.

Verfaellie, M. (2000). Semantic learning in amnesia. In L. S. Cermak (Ed.), *Handbook of neuropsychology* (vol. 2, pp. 335–354). Elsevier Science.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning. *Neuron*, *31*(2), 329–338.

Ween, J. E., Verfaellie, M., & Alexander, M. P. (1996). Verbal memory function in mild aphasia. *Neurology*, *47*(3), 795–801.

Westmacott, R., Black, S. E., Freedman, M., & Moscovitch, M. (2004). The contribution of autobiographical significance to semantic memory: Evidence from Alzheimer's disease, semantic dementia, and amnesia. *Neuropsychologia*, *42*(1), 25–48.

Westmacott, R., & Moscovitch, M. (2003). The contribution of autobiographical significance to semantic memory. *Memory and Cognition*, *31*(5), 761–774.

Wickelgren, W. A. (1965). Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 53–61.

Wiggs, C. L., Weisberg, J., & Martin, A. (1998). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia*, *37*(1), 103–118.

Wiskott, L., & Sejnowski, T. J. (2002). Slow feature analysis: Unsupervised learning of invariances. *Neural Computation*, *14*(4), 715–770.

Yassa, M. A., & Reagh, Z. M. (2013). Competitive trace theory: A role for the hippocampus in contextual interference during retrieval. *Frontiers in Behavioral Neuroscience*, *7*, 107.

Zito, T., Wilbert, N., Wiskott, L., & Berkes, P. (2008). Modular toolkit for data processing (MDP): A Python data processing framework. *Frontiers in Neuroinformatics*, *2*, 8.

_____