INVESTIGATING THE ROLE OF PREDICTIVE REPRESENTATIONS IN IMPLICIT EVENT BOUNDARIES, STATISTICAL LEARNING, AND CATEGORIZATION.

A Dissertation Presented

by

TEJAS SAVALIA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2024

Psychological and Brain Sciences

INVESTIGATING THE ROLE OF PREDICTIVE REPRESENTATIONS IN IMPLICIT EVENT BOUNDARIES, STATISTICAL LEARNING, AND CATEGORIZATION.

A Dissertation Presented

by

TEJAS SAVALIA

Approved as to style and content by:

Andrew Cohen, Chair

Jeffrey Starns, Member

Youngbin Kwak, Member

Meghan Huber, Member

Maureen Perry-Jenkins, Chair
Psychological and Brain Sciences

DEDICATION

To those little lost sheep.

ACKNOWLEDGMENTS

Thanks to all those fine shepherds. Not to mention all the great border collies and suchlike fine animals.

ABSTRACT

INVESTIGATING THE ROLE OF PREDICTIVE REPRESENTATIONS IN IMPLICIT EVENT BOUNDARIES, STATISTICAL LEARNING, AND CATEGORIZATION.

SEPTEMBER 2024

TEJAS SAVALIA

B.E., GUJARAT TECHNOLOGICAL UNIVERSITY

M.S., INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY, HYDERABAD

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Andrew Cohen

Sheep like grass. Why? Let me tell you. Sheep are ruminants, like cattle, deer, and horses. They have stomachs that are specialized...

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

We experience a constant stream of sensory information from the moment we are born. Our brain parses this information and slowly learns to extract meaning from it. From recognizing the mother's scent as a survival instinct to formulating complex plans to defeat a board game opponent, our brain extracts meaning from our surroundings, considers prior experience and the current state of the world, and makes decisions to interact accordingly. My key question in this dissertation is: How do we learn to extract meaning from our surroundings?

Extracting meaningful information from our experience is beneficial to our functioning and survival. When approached by a cheetah in a forest, we do not wait to evaluate the exact number of spots on its skin before deciding to run. Such abstractions and formation of heuristics allow us to process naturally complex surroundings in quick time and act accordingly.

When extracting meaning from our surroundings, we often (need to) ignore minute details and abstract out towards a coherent thought of our surroundings. The extreme example above aside, we observe such abstractions in almost all day-to-day activities. Imagine someone asking you about the events during the day before reading this manuscript. Perhaps you are reading it on your office computer and you recall a sequence of events starting from waking up, to getting ready, driving or taking public transit to work, getting your morning coffee and breakfast in with an email check before turning your attention to this manuscript. Each event described above combines several sub-events that are abstracted away in such a verbal recall and description. For example, getting ready involves several steps from brushing your teeth, showering, and wearing your work outfits. Each sub-event can be further thought of as an abstraction from sub-sub-events – brushing your teeth is a combination of putting toothpaste on the brush head, the physical act of brushing, followed by rinsing. While we perform each act continuously in time, our recall (and by extension, representation in memory) of these past events is discrete, segmented, and abstracted. The meaningfully separatable chunks of a continuous stream of events help in storage in (and retrieval from) memory.

There is often agreement on what it means for a chunk and for boundaries defining such chunks to be 'meaningful'. In the example above, it is reasonable to argue that brushing teeth, showering, and putting on clothes are three distinct activities. Furthermore, even when the true transitions between these activities are continuous and seamless to a independent, naive observer, the boundaries between these events are perceptually meaningful. What aspect of the environment dictates this agreement about the points at which we segment events and what is special about the properties of the events between those points that lead to distinct representations in memory?

One could argue that these events that occur at different points in time also occur at different spatial locations, thus providing different contexts and hence separate representations in our brain to be considered distinct. Transitioning from one (temporal) event to another can thus be akin to transitioning from one room of the house to another. However, it is almost impossible to decouple temporal events from spatial events assuming a causality from spatial segmentation to temporal segmentation for any spatially experienced distinct event is also a temporally experienced distinct event. Instead, arguing that temporally distinct events lead to a spatially distinct representation can provide a more encompassing explanation of distinct representations for events in distinct spatial *and* temporal contexts. One could similarly argue that the formation of "meaningful" chunks is through perceptual differences between the events we experience. While lower-level perceptual experiences are indeed often different for different events, the mapping of perceptual differences onto segmented events is arbitrary and not a *sufficient* condition. For example, eating an apple is recalled as eating an apple regardless of whether it has a green leaf added to its top. Perceptual distinctions are not enough to determine whether events are represented distinctly. What then is the key mechanism that leads to events being segmented?

In this dissertation, I argue that the primary reason and mechanism through which we segment events is based on temporal contingencies of various sub-events that encompass an event. Specifically, we recall being in the kitchen as different from being in the bedroom because we have a coherent set of experiences in the kitchen that are distinct from a coherent set of experiences in the bedroom. For an infant forming knowledge of the world, a kitchen while perceptually distinct from a bedroom, is not meaningfully different. With experience and observation of the functions within these spatially (and perceptually) distinct locations, the child slowly develops distinct representations of the two rooms.

This dissertation focuses on understanding temporal contingencies' role in event segmentation, and by extension, general pattern extraction. I argue that even *without* any spatial or perceptual information that may aid us in separating events in memory, we can use temporal coherence to experience separate events and abstract information to aid higher-level cognition. Specifically, I investigate the parsing of a continuous stream of information into discrete chunks in three ways:

- The possible algorithmic processes that naturally lead to such segmentation and the impact of environmental properties in aiding this abstraction.
- The properties of the temporal boundaries when such temporal abstraction occurs naturally and implicitly.

- The role of temporal events separated by implicit temporal boundaries in forming abstractions.

1.1 Scope of this dissertation

The human brain is a complex machine – millions of neurons act as computational units and combine in specific ways to form a functioning human being. These neurons come together to implement several levels of function from lower-level automatic perception to higher-level planning and conscious thought. This dissertation does **not** focus on these implementational-level mechanisms of cognition. Rather, it focuses on *algorithmic* computations that neurons may, collectively, implement that lead to us acquiring patterns in the environment around us.

Analyses in this dissertation use several models of cognition in investigating the role of implicit statistical learning. In most cases, the focus of this dissertation is **not** to evaluate the validity of these models. Indeed, most models of cognition are wrong; but are useful and I use several such models to evaluate specific aspects of how we acquire patterns.

1.2 Format of this dissertation

In the rest of this dissertation, I present three lines of studies investigating the role of implicit temporal boundaries in cognition. In Chapter 2, I present an algorithmic framework that naturally leads to a representation of separable events without the need to rely on explicit properties of the experienced events. I also show that this framework allows for a distinct representation of event boundaries as special events. In Chapter 3, I present work comparing the properties of these event boundaries which are operationalized implicitly (i.e. through no perceptually special information) with event boundaries as they have been studied in prior literature which are operationalized explicitly. Finally, in Chapter 4, I present work investigating the role of these implicitly operationalized event boundaries in categorization to serve as a gateway for understanding higher-order cognition in the context of temporal segmentation and pattern acquisition.earned.

CHAPTER 2

QUALITY OF ENVIRONMENTAL EXPOSURE MODULATES STATISTICAL LEARNING

2.1 Introduction

Imagine you just moved to the United States and are visiting Target for the first time. Perhaps since you just moved in, your first goal is to furnish your apartment. You look around at the entrance, and navigate your way to the furniture section perhaps while taking a few false turns on the way, buy the stuff you need, pay and leave. The next time you visit for, for example, groceries and produce. You visit again for sporting goods and then again for gifts for your friends. An year later, all such subset of visits through the Target store makes you an expert in knowing the specific route to the section you need to visit. What algorithmic mechanisms allow us to build such expertise to create connections in an explored environment even when during no single visit, you explored all possible connections between regions of the store?

We are able to build up a map of the environment without being exposed to the full extent of it in a single iteration simply based on local exposures to it. Such building is often implicit – you just know that in order to go to the furniture area, you need to pass through the gifts even when you did not explicitly explore this connection before. In this chapter I explore some algorithmic mechanisms that may lead to our ability to acquire large-scale structural knowledge of the environment based on local exposures.

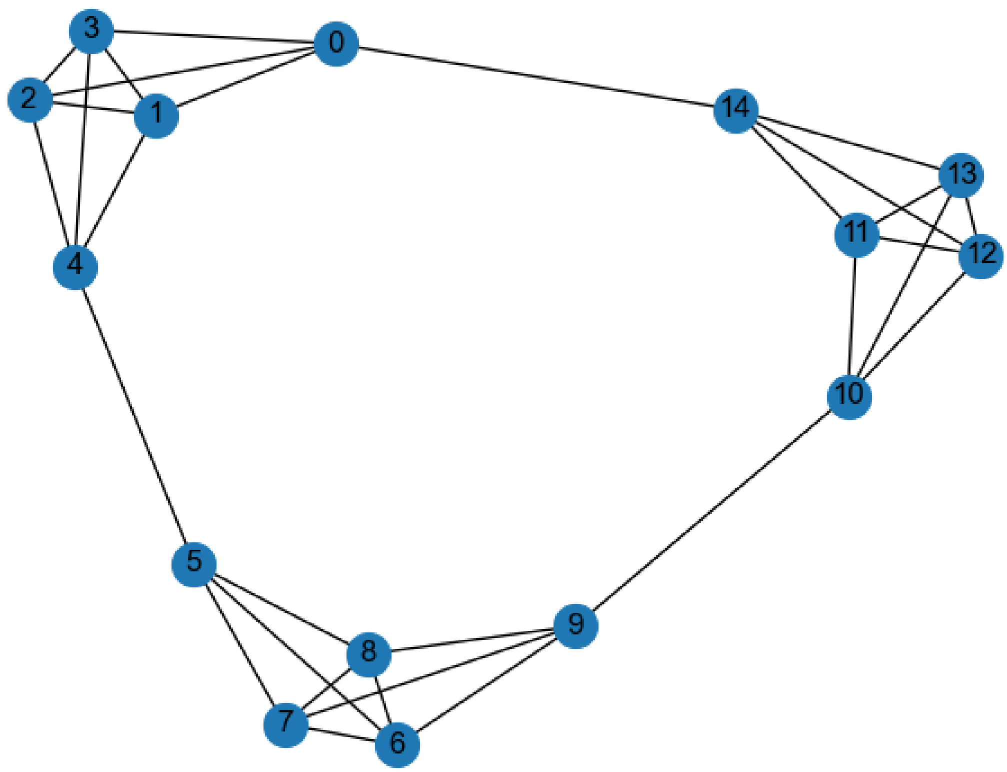
The effect of local exposure in acquiring structural knowledge of the environment have been explored in several areas of cognitive psychology through artificial grammar and language learning [38, 66, 1, 17], visual statistical learning [24, 81, 8], or motor sequence learning [3, 54, 12, 34]. In recent work, (implicit) acquisition of higher order knowledge of the environment from lower order exposure is studied through structured graph based transitions between stimuli [71, 37, 34, 45, 47, 35]. For example [71] show that when asked to respond to arbitrary stimuli arranged in a temporally graph-modular structure, participants often parse the edges that connect two modules as 'natural breaks' even when their local exposure does not distinguish between the cross-module and within-module edges.

More commonly, global-scale structure acquisition has been attributed to response time measurements. Earlier work in serial reaction time tasks [54, 12] shows that breaking an implicitly learned motor sequence leads to slower reaction times. The slowed reaction times when crossing the between-module edges have also been shown in recent work on statistical learning in modular graph structures [34, 45, 37, 35, 36, 47].

This slowdown across module edges appears to be mediated by the nature of the walk experienced across the community structure where random and Eulerian walk (a walk where each edge of the graph is visited exactly once before repeats) experiences continue to show this slowdown whereas a Hamiltonian walk (a walk were each node of the graph is visited exactly once before repeats) experience does not [37]. Thus it appears that the kind of experience through the graph alters the knowledge of underlying statistical patterns. Similarly, the topographical structure of a graph in motor skill learning tasks also appears to alter structural knowledge [46, 47, 45] where modular graphs like in Figure 2.1 produce the largest dip in reaction times when responding to boundary items.

Why do we slow down at boundary nodes that lead to the adjacent module even when the local probability of that particular transition is the same as any other transitions? Understanding this particular property of human behavior may provide

Figure 2.1. Modular graph structure used in [71]. Locally, each node is connected to four nodes with each edge equally probable. However, globally, the graph structure consists of three sub-modules interconnected through 'boundary nodes'



deeper insights into what leads to global-scale structure acquisition; after all the only difference between the boundary node and other non-boundary nodes is in context of the global structure of the graph. Event boundary literature (where boundaries are typically operationalized through explicit changes in context) suggests that boundaries alter the predictability of future events and this predictability leads to event segmentation [89, 13]. Thus, in implicitly operationalized boundaries such as in serial reaction time tasks, the slowdown at the boundary node may imply a similarly increased uncertainty at boundary nodes leading to slowed responses. Prior work aimed at understanding human representation of graph structures indeed points to an the ‘cross-entropy’ between a learner’s estimate of the transition probability and the true transition probability of the environment [46, 45, 47].

In particular, [47] show that algorithms of contextual representations such as the Successor Representation (SR) model in Reinforcement Learning [16, 53, 26] or the associative learning based Temporal Context Model (TCM) can naturally lead to an increased cross-entropy for cross cluster transitions relative to within cluster transitions in modular graphs. In this work, by using the framework of cross-entropy to estimate reaction times in a modular graph we aim to 1) Experimentally test the predictions of these two models when exposure through the modular graph structure of is partial and 2) Provide evidence in favor of one of the two models of representation.

2.1.1 Representations of Temporal Context

2.1.1.1 Successor Representation

The Successor Representation (SR) model of reinforcement learning has been used as a model to understand the generalization of reinforcement learning behavior in large action spaces [16]. In recent work, the SR model has also been shown to be a reliable model for explaining human decision-making behavior in multi-step environments. The model accurately predicted that humans are worse at adapting to changes in the transition probability of a learned environment than to changes in the end-point rewards [53]. There has been further evidence of SR being represented in the Hippocampal cells which represent space [26, 76].

Briefly, the SR model represents each state in the actionable space as a predictive representation matrix. For an environment of N discrete states, the SR matrix M of size (NXN) maintains expected future visits to a given state from each state. Specifically, element $M_{i,j}$ of the matrix represents the expected future visits to state j from state i. This transition matrix is learned over time based on the temporal difference error learning rule. For example, consider at a given point in time, t, an agent maintaining the SR matrix is in state i. The agent now moves to state j out of the possible N states. The i^{th} row of the SR matrix is updated as follows:

$$M_{i,j} \leftarrow \hat{M}_{i,j} + \alpha[\delta(s_{t+1}, j) + \gamma * \hat{M}_{s_{t+1},j} - \hat{M}_{s_t,j}] \tag{2.1}$$

where $\delta(.,.)$ equates to 1 if both arguments are equal otherwise it equates to 0. Thus, the matrix increases the probability of visiting a state j from state i if state j is visited in the current experience and it decreases the probability of visiting all other states from state i. Parameter α is a learning rate parameter that determines how much of the previous estimate of visiting state j from i is factored into the current update. Parameter γ is a future discount parameter that dictates how much in the future the agent sees – specifically, a higher value of γ indicates future visitations to state j are weighed high in the current update.

2.1.1.2 Temporal Context Model

The Temporal Context Model (TCM) was devised to explain the primacy and recency effects in human recall and recognition memory [33]. The TCM model assumes that the items maintain a temporal context as they get encoded thus allowing items presented close to the previous items to share such temporal context. Briefly, the TCM can be formalized as [27]:

$$\begin{gathered} t_n = \rho * t_{n-1} + f_n \\ \hat{M}_{i,j} \leftarrow \hat{M}_{i,j} + \alpha f_{n+1} t_{n,i} \end{gathered} \tag{2.2}$$

where t_n is said to be a 'context' vector for item n. The context drift parameter ρ determines the proportion of the previous elements's context that gets incorporated in the current context. f_n is a one-hot encoded vector for item n. The learning rate parameter α determines what proportion of the currently experienced state binds with the existing context.

The key difference between the two models of temporal context is two fold: (1) SR Relies on error-based learning whereas TCM relies on hebbian, assosciative learning and (2) Through the future discount parameter γ, SR also learns the predictability observing states in the near future based on the locally experienced transitions [27].

2.1.2 Model Simulations

The models described above can be used to simulate expected behavior as a function of a range of exposure. Figure 2.1.2 shows the context matrix representation after the models have been simulated for a random walk through the graph structure in 2.1 as a result of a random walk after 1000 trials for both models. Previous work has shown that participants acquire the global structure of the graph for a random walk.

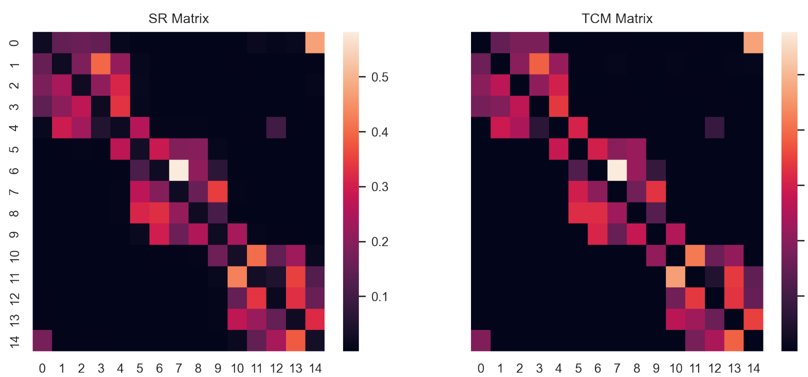


Figure 2.2. Successor Representation and Temporal Context Model representations of context following a random walk through the modular graph structure.

To model the observed differences in reaction times and link them to the apparent differences shown in Figure 2.1.2, we apply principles of information theory. Specifically, we assume that response time for each stimulus is a function of the uncertainty in its surrounding context. Measures of information entropy have previously been used to explain RT differences between cluster transitions while traversing similar graph structures [46, 47, 45]. Formally,

$$RT(node) \cong Entropy(node) = \sum_{s' \in S} \hat{M}(s,s') * log(\hat{M}(s,s')) \tag{2.3}$$

where $M(s,s')$ is the context representation at node s. For SR, this expression evaluates to the expected future visits to state s' from state s whereas for TCM this expression evaluates to the extent to which s' is activated as a result of s.

As noted previously, a common indicator of participants having acquired the global structural knowledge is a slow down in responses when the ongoing stimulus stream crosses a cluster (relative to transitions within a cluster) of the modular graph. Context representations can be used to model the cross cluster transitions by computing a 'surprisal' effect. For simulations, the surprisal effect is computed as the Jenson-Shannon distance between the context representations of two nodes. Formally,

$$RT(s \rightarrow s') \cong JS(s,s') = \sqrt[2]{\frac{D(M(s,.)||p) + D(M(s',.)||p)}{2}} \tag{2.4}$$

where $M(s,.)$ is the context representation of node vector s, p is the point-wise mean of nodes s and s' and $D(M||p)$ is the Kullback-Leibler divergence between probability distributions M and p.

The formalization of observed response time differences due to surprisal (and node entropy) allows us to simulate expected reaction time distributions for novel walk types. Specifically, to understand the mechanisms behind acquiring the global modular graph pattern following a limited exposure, each model was simulated for random walk with lengths of 0, 3, 6, and 999. A random walk length of 0 translates to a completely random selection of one of the 15 nodes of the modular graph on each trial. Walk length of 3 and 6 translate to a random walk visiting 3 and 6 edges (4 and 7 nodes) respectively before being reset to a random node (similar to visiting the Target store in short bursts to purchase relevant items and checking out without visiting the entire store). Finally, a walk length of 999 translates to visiting 999 edges (with repetition) through their connections on the modular graph.

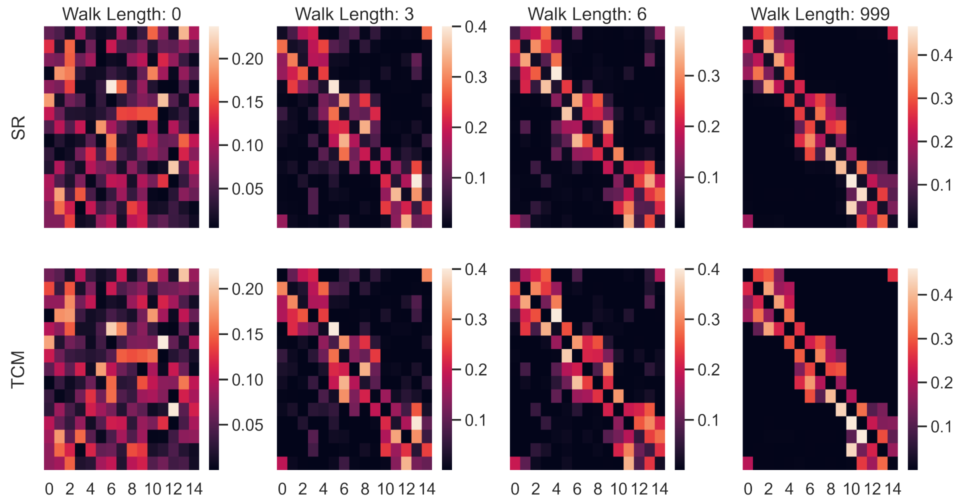


Figure 2.3. Model prediction of context representations for SR and TCM models across different walk lengths. Both models seemingly predict that the modular structure of the original graph is increasingly recovered with longer walk lengths.

Parameters of the simulations in figure 2.1.2 are determined through a best-fitting grid search procedure. Specifically, for a combination of parameters the euclidean distance between the genearted context matrix (SR or TCM) was computed. A grid search was conducted to determine the parameters which minimized this euclidean distance.

The acquisition of the global structure can be modeled using surprisal as has been done in previous research [46, 45, 47]. For a subset of parameters in the valid range of 0 to 1, each model was simulated to produce a context matrix. Jensen-Shannon distance was computed between each pairs of nodes and averaged over cross-cluster pair and within cluster pairs. Simulation results below show the transition Jensen-Shannon distances over 100 simulations of the model for each parameter combination. For SR, 'param_a' is the learning rate parameter α and 'param_b' is the discount parameter γ. For TCM, 'param_a' is the learning rate parameter α and 'param_b' is the context drift parameter ρ.

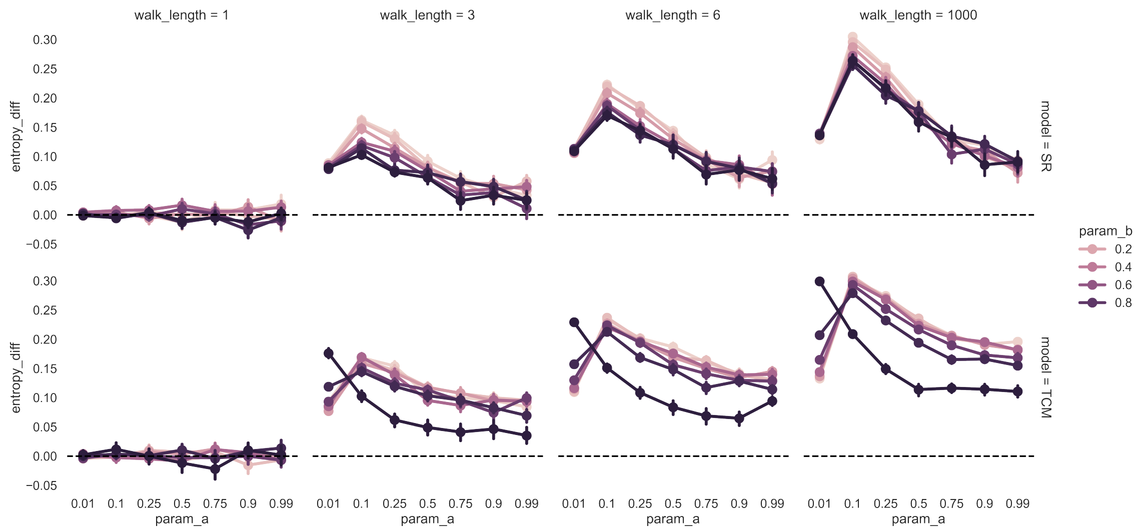


Figure 2.4. Model Predictions for differences in reaction time comparing across cluster transitions to within cluster transitions across walk lengths. Both models predict that cross cluster surprisal effect will increase with walk length leading to an increased reaction time.

Figure 2.1.2 shows that both context models predict an increased surprisal as walk length through the modular graph gets longer. As walk length increases, context associated with each node increasingly represents neighboring nodes. Since neighbors of the boundary nodes differ more than those between the non-boundary nodes, crossing from a boundary node to another boundary nodes.

The two context models, however, differ in their predictions in the role of a boundary node. Figure 2.1.2 shows that SR predicts an increased entropy in its representation of the boundary nodes with walk length relative to the non-boundary nodes for some values of the α and γ parameters. On the other hand the TCM does not predict such increased in Boundary vs Non Boundary entropy differences.

The predictive nature of SR (as modeled by the future discount, γ parameter) allows for a representation of nodes in the neighboring cluster to impact entropy on the boundary node of the current cluster that leads to that neighboring cluster. This effect is unique on boundary nodes of a cluster as non-boundary nodes of the second

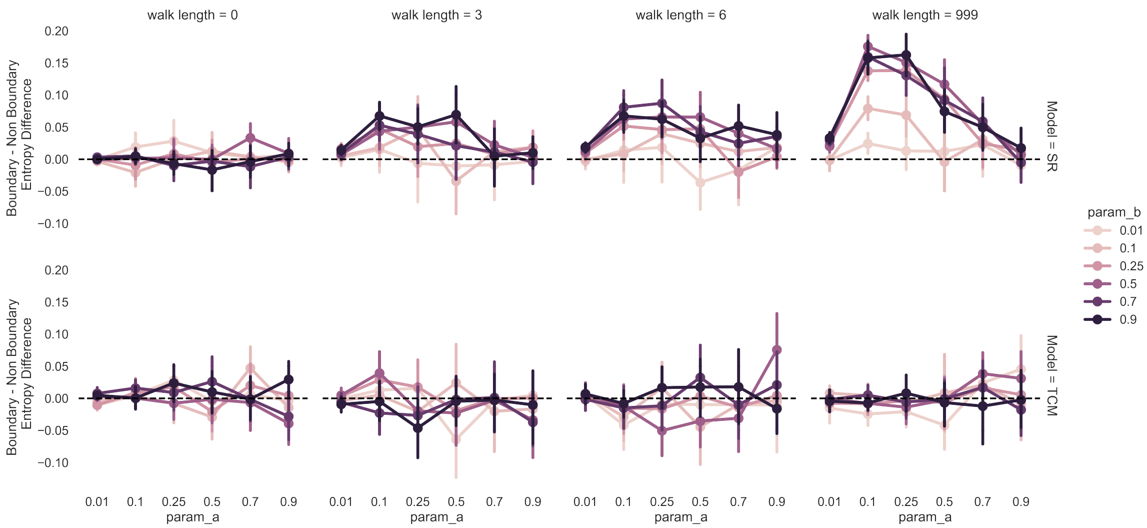


Figure 2.5. Model prediction differences between the SR and the TCM after different walk lengths. SR predicts that entropy of boundary nodes will scale with walk lengths whereas TCM does not.

cluster are closer to the immediate neighbor of the current cluster (i.e. the boundary node that serves as an entry point to the second cluster) Since TCM is associative (as opposed to predictive), each node activates its immediate and two-step neighbors with equal weight. associative nature of TCM does not transfer representations of nodes of the neighboring cluster to the boundary node of the current cluster. Rescaled heatmap in figure 2.1.2 presents this effect.

The SR-based predictive context representation in particular shows that boundary nodes carry more information than non-boundary nodes whereas the associative context representation does not produce this effect. [1]

Thus, both SR and TCM models would predict slow down in cross-cluster transitions relative to within cluster transitions, and that this slow down will increase with walk length. However predictive context representations through SR is unique

[1]The activity in the lower third of both matrices is due to recency; while these are interesting patterns, and seem to indicate that SR can account for the recency effects in memory which was the primary motivation behind introduction of the TCM [27, 33]. Investigating recency effects in this implicit statistical learning context is out of scope for this dissertation.

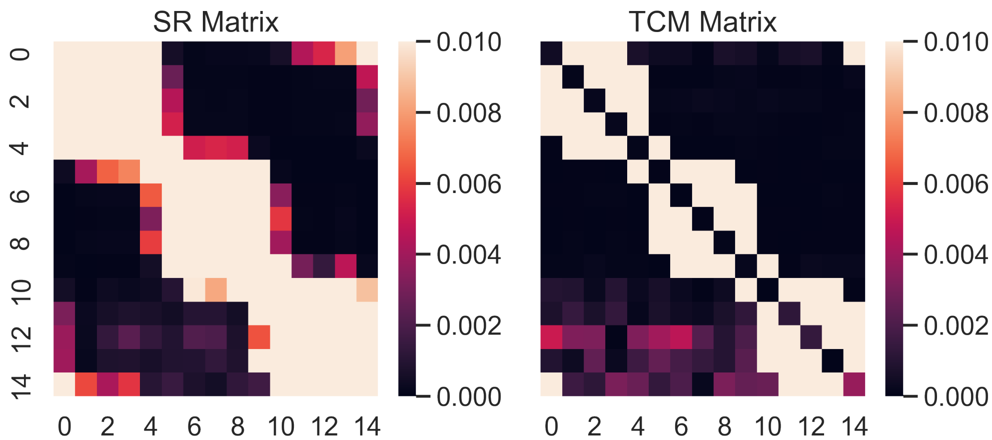


Figure 2.6. Rescaled SR and TCM matrices depict differences between context representations of the two models. Boundaries in SR incorporate more information than those in TCM.

in predicting the scaled slow down at boundary nodes with random walk length, *independent* of transitions. While lack of a scaled slow down to boundary nodes does not invalidate the SR model (because some values of the parameters allows SR to not scale the slowed reactions with walk length), the presence of such a slow down provides evidence for predictive representations in such statistical learning tasks. The study presented next, thus tests this prediction.

CHAPTER 3

COMPARING IMPLICIT EVENT BOUNDARIES WITH EXPLICIT EVENT BOUNDARIES

3.1 Introduction

We receive a continuous stream of sensory information in our daily lives. In order to make sense of it, we often parse it into meaningful chunks for storage, retrieval and comprehension. For example, we may recall our drive to work as a series of discrete events; got into the car, got coffee, picked up a collegue, hit traffic on a particular street, parked, and walked over to the office. What aspects of the incoming stream help us organize continuous temporal information in such discrete chunks? Temporal chunking in cognitive psychology has been studied under several domains from event boundaries [13, 89, 68, 67, 21, 3], language learning, [66, 38], categorization [82, 25], and motor sequencing [5, 80, 70, 59]. Chunking a repeated sequence of experiences is crucial to abstracting patterns in the environment and formation of habits for quick and efficient interactions with the environment [18, 74, 20, 19, 28, 7].

Models of temporal event segmentation suggest that the points which lead to temporal segmentation seem to be unique in their properties in both segmenting the continuous stream of information and integration of information across the temporal event. These ‘event boundaries’ are, for example, shown to be remembered better [79, 67, 67, 87, 60, 31], serve as points of retrieval [51] and replay to promote long term memory [30, 75] and easy parsing, help integrate memory across time [13], and separates across boundary events while collapsing within boundary events [13, 42, 23, 9].

In most prior studies, event boundaries have been studied using explicit context shifts. For example, when stream of stimuli are surrounded by colored border, event boundaries are operationalized by first showing the stimuli surrounded by a color and abruptly changing that color[31]. In another study, event boundaries were operationalized via explicit context changes by changing the associated stimulus[23]. A pair of images were presented on each trial; one image of the pair, the 'scene' image remained constant for a short sequence of trials whereas the other ('object' or 'face') changed on each trial. Participants were asked to make judgments about the object/face image [23]. Previously, context changes had been operationalized as either perceptual or semantic shift in ongoing set of events by having participants watch clips [79]. In more recent work, context change has been operationalized as changes in ongoing reward contingencies associated with each stimulus [68].

Consistent findings across most studies in explicitly operationalized event boundaries show that event boundaries are often remembered better [79, 60, 31, 13, 68, 23, 2], and events across boundaries appear to be perceptually farther whereas events within boundaries appear to be perceptually closer [13, 23, 9, 42]. In recent work, however, it has been shown that event boundaries can also be formed *without* explicit changes in context. After being exposed to a stream of stimuli such that the ordering is controlled by a modular graph shown in figure 2.1, participants seem to recognize across cluster transitions as 'natural breaks' more often than within cluster transitions [71]. In recent work, this finding has been linked to statistical learning of temporal graph structures [35, 36, 34, 34, 46, 47, 45] and measured by slowed reaction times across clusters than within clusters. However, past studies where boundaries are operationalized implicitly do not assess the memory representations of these boundaries using the same tests used in explicitly operationalized boundary paradigms.

In this chapter, I present two tests on implicitly operationalized boundaries to assess whether they elicit the same behavioral properties as the explicitly operationalized boundaries. In particular, I use the paradigm and graph structure previously used in Schapiro et al. [71] to test whether participants recall boundary items better (or worse) than non-boundary items. I then use a two module graph structure in Figure 3.3 to test whether items across the two clusters appear farther than items within a cluster (similar to findings in explicitly operationalized boundary paradigms).

3.2 Modeling Boundary Memory Benefits

Event segmentation theory suggests that the segmentation of the continuous sensory experience occurs automatically and through prediction errors[89, 88, 79]. According to the event segmentation theory, we maintain an ongoing 'context' which is predictive of upcoming events. Event boundaries are created when this prediction breaks. More recent work has shown that prediction errors are not necessary for creation of event boundaries; a change in uncertainty of the upcoming events can also produce event boundaries [73]. Prediction errors particularly lose their value in learning new information when the explored environment is uncertain[4]. Nevertheless, under environments with high regularities, prediction errors remain the key mechanisms driving boundary formation.

As reviewed above, prediction errors need not be explicitly operationalized for an event boundary to be learned. Prediction errors which imply shifts in ongoing context, similar to implicitly operationalized event boundaries may be implicit. In chapter 2 show that context models can be used to estimate representations of implicitly operationalized event boundaries. Particularly, predictive representations such as the SR provide a natural representation of event boundaries which form bottlenecks in transitioning between clusters in modular graphs in figure 2.1. I propose that the same predictive context-representation framework using the Successor Representation model of Reinforcement Learning [16, 53, 69, 52, 27] can be used to model differences in memory representations.

To simulate a recognition memory task, I employ a simplified version of the exemplar-based Generalized Context Model [56, 55, 58]. The GCM follows a class of global matching exemplar models where each studied item is stored as an image or an exemplar in memory. At test, the presented test item is matched with memory representations of stored exemplars by computing the psychological similarity between them. It is assumed that if the similarity, summed over all similarities of the test items with exemplars in memory, crosses a criterion (a free parameter), the participant recognizes that item and the 'old' response is chosen in the old/new recognition test. Similarly, a 'new' response is chosen when the summed similarity of the test item with all stored exemplars falls below the decision criterion.

The GCM model for recognition memory can be formalized with the following equations [58]:

$$\begin{aligned} d_{ij} &= [\sum_{k=1}^{K} w_k (x_{ik} - x_{jk})^2]^{1/2} \\ s_{ij} &= \exp^{-c_j d_{ij}} \\ a_{ij} &= m_j s_{ij} \end{aligned} \tag{3.1}$$

where d_{ij} is the psychological distance between test item i and examplar j, w_k is the weight a participant may place on the k^{th} dimensions (and $k \in K$). The distance metric is thus computed as an euclidean distance between exemplars in memory and the test item weighted by where each feature is allowed to have a different weight to reflect differentially important features. sij is the similarity between test item i and exemplar j which decreases exponentially with psychological distance. c_j is a scaling factor determining how much the similarity falls off for a unit of distance for each exemplar. a_{ij} is the activation of exemplar j when compared with test item i and is scaled by the memory strength of the exemplar m_j.

To demonstrate the potential role of temporal structure, a few simplifying assumptions are made to the recognition memory model. Specifically, in simulations presented below, it is assumed that the each feature dimension of the studied (and test) items is weighted equally. Furthermore, the scaling parameter c is assumed to be 1 for all exemplars. To simulate the differences between boundary and non-boundary nodes in memory, it is assumed that the memory strength of an item associated with each node is proportional to the entropy in its successor representation of that node. While evidence for relating memory strength to context based entropy is scarce, past work has shown that entropy (as a measure of uncertainty) has been a helpful factor in motivated learning and is a contributing factor in hippocampal activation [15]. Furthermore, the slow down associated with increased entropy as demonstrated in previous statistical learning tasks [46, 47, 45] implies that participants at the least spend more time on such high-entropy boundary nodes, thereby allowing for a better chance of remembering these nodes better.

Given these assumptions, simulating recognition memory on the final SR representation provides an expected comparison of recognition memory acccuracy for old boundary, old non-boundary and new items. Figure 3.2 shows the what this modeling approach expects. On average, implicitly operationalized boundaries are expected to be remembered better than non-boundaries. This benefit is expected to be apparent for participants who are exposed to the temporal structure (right panel) relative to participants who are not (left panel).

3.2.1 Methods

3.2.2 Results

3.2.2.1 Diffusion Modeling

Preliminary results above reflect that while there are no differences in memorability of boundary items, non-boundary items on the other hand tend to be remembered worse for participants in the structured conditions relative to the structured conditions. However, the model used above fails to account for ceiling effects – accuracy

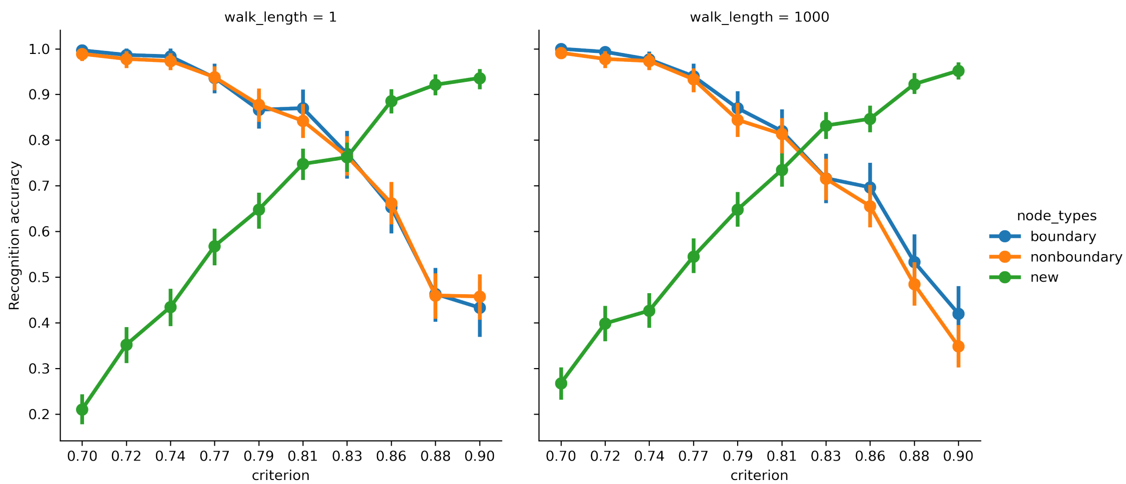


Figure 3.1. Simulated recognition memory test performances for walk lengths of 1 and walk lengths of 1000 on modular graph in Figure 2.1. On average, recognition memory performance is expected to be better for boundary items than non-boundary items.

for old items is near perfect or could have reached an asymptote. Those models also do not account for encoding accuracy during exposure – participants may fail to recognize old items if they did not study those items well as indicated by the rotation judgement task.

To be able to account for both ceiling effects and effects of exposure, we can use additional information available in form of response times during the recognition memory task. For example, for participants equally accurate in recognizing boundary and non-boundary participants, being able to recognize boundary items faster may provide additional evidence for better memorability of these items relative to slower recognized non-boundary items. Figure **??** shows median response times across three blocks of recognition test.

To understand recognition memory differences between boundary and non-boundary items in context of response accuracy and response time distributions, we use the Drift Diffusion Model DDM. The DDM, which falls under a class of sequential sampling models, has been a widely successful model in modeling two-choice tasks in recognition memory [63, 61, 77, 78, 62]. The function of the DDM is depicted in Figure 3.2.2.1.

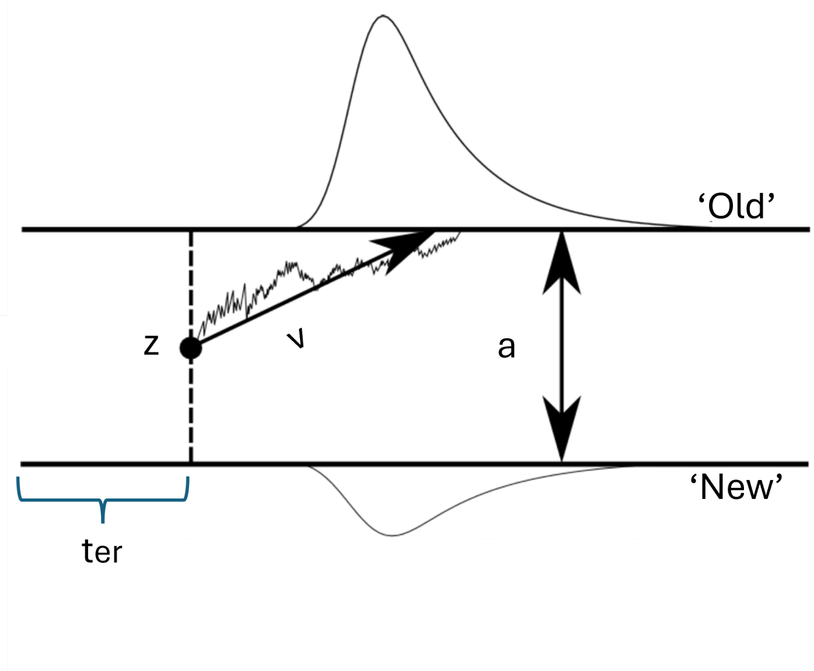


Figure 3.2. The Drift Diffusion Model of Choice Response Times for Old/New recognition memory tasks.

Briefly, the DDM assumes that at choice time, evidence from two presented options accumulates sequentially over time towards one of the two boundaries. For a previously studied item presented at test, the evidence from the item accumulates slowly towards the 'old' boundary whereas for a non-studied item, the evidence accumulates towards the 'new' boundary. The rate of evidence accumulation is controlled by the drift rate parameter, 'v'. Participants may be biased towards making an old or a new response at test; this bias is measured by the starting point parameter, 'z'. THe boundary separation between the two responses is modeled by a parameter 'a'. Finally, the observed response consists of cognitive processes not affiliated with decision making (such as time it takes to visually process the test item, time for the motor systems to click the relevant key) which are modeled by a non-decision time parameter 't_{er}'. [1]

Prior work has shown that memory strength of previously studied items impacts the drift rate towards old/new responses. The drift rate parameter implies that a stronger match to memory leads to a quicker accumulation of evidence towards the 'old' response boundary. Similarly, a stronger mis-match to memory allows for a quicker accumulation of evidence towards the 'new' response boundary[63, 61].

The DDM thus allows us circumvent ceiling effects by modeling response time distributions and estimate whether if boundary items are indeed a remembered better than non-boundary items in the structured exposure or whether the effect is driven by worse-remembered non-boundary items.

3.3 Modeling Boundary Distance Effects

Another replicated finding in explicitly operationalized event boundary literature is an apparent increased separation of events across boundaries [32, 9, 21, 22, 31]. This separation of events across boundaries helps shape narratives in long term memory [13].

It is unknown, however, whether the increased temporal separation across event boundaries generalizes to boundaries operationalized implicitly as well. Context models such as SR provide a mechanism to directly estimate perceived distance between events. Specifically, each cell in the SR matrix $M(s, s')$ indicates the future expected visit probabilities from state s to state s'. Under the assumption that states closer to the current state are visited more often in a random walk than state farther, the probability $M(s, s')$ provides a direct estimate of perceived distance from node s to s'.

[1] Note that this version of the DDM is a simplified model. More complex DDMs account for trial-to-trial variability in each parameters as well.

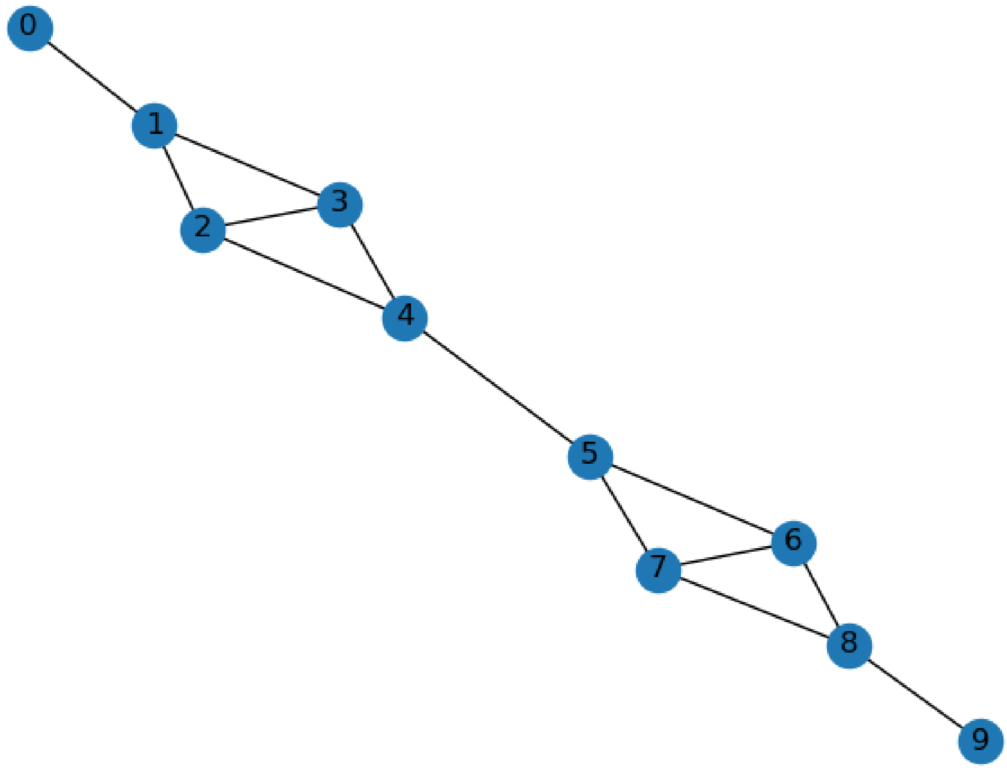


Figure 3.3. Two module graph used for distance judgments.

To test whether cross boundary events are perceived to be farther from each other relative to within-boundary events graph in figure 3.3 is used. The key transitions of interest that are compared in this graph are transitions at equal distances. Below I specify example transitions at 3 different distances however, for simulations and the experiment, all symmetrical transitions at those distances were tested.

- $0 \leftrightarrow 4$ vs $1 \leftrightarrow 5$ at shortest distance 3.
- $1 \leftrightarrow 4$ vs $2 \leftrightarrow 5$ at shortest distance 2.
- $2 \leftrightarrow 4$ vs $4 \leftrightarrow 4$ at shortest distance 1.

SR estimate of distances is shown in figure 3.3. Simulations predict that while boundary nodes themselves get perceptually farther with increased discount rate,

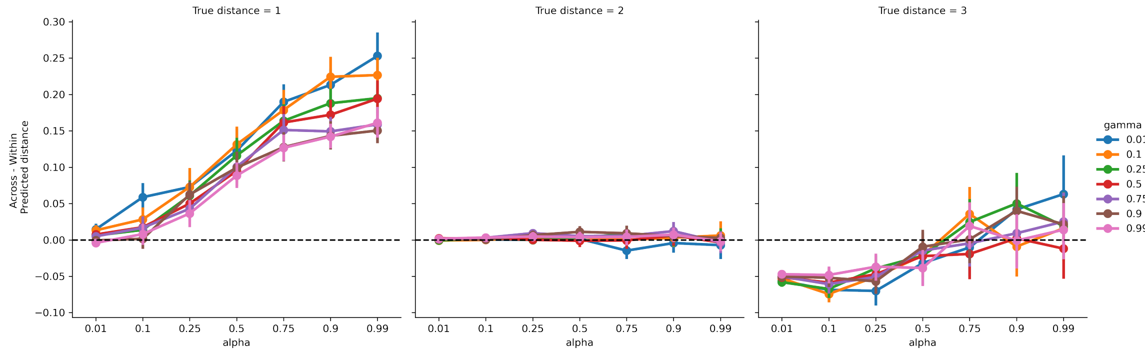


Figure 3.4. SR predictions of distances across boundaries relative to distances within boundaries for nodes at true distance of 1, 2, and 3 and different parameter combinations.

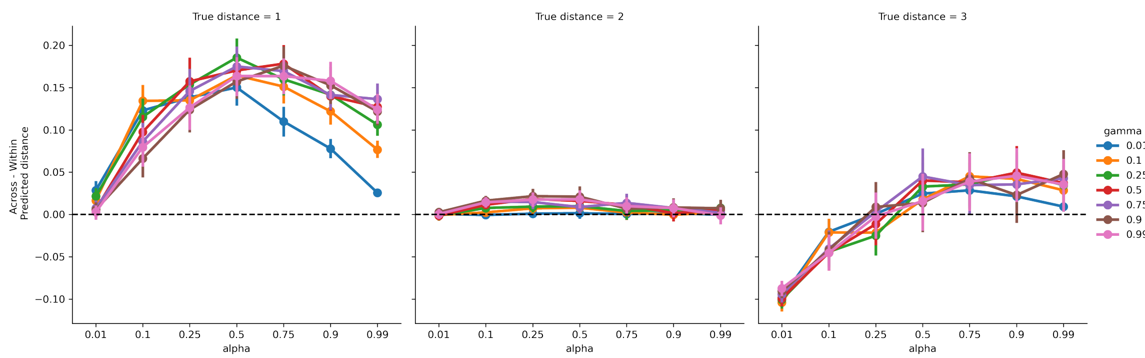


Figure 3.5. SR predictions of distances across boundaries when boosted by node entropies relative to distances within boundaries for nodes at true distance of 1, 2, and 3 and different parameter combinations.

neither of node-pairs at distance 2 or 3 that involve boundary nodes reliably show an increased cross-cluster distance.

However, in a typical experience of the random walk, participants are shown to be slow at responding to boundary nodes; both due to higher entropy at boundary nodes (from chapter 2) and previously shown increased response times during transitions across temporal clusters [46, 34]. On average, a participant would typically spend more time at boundary nodes than non-boundary nodes thereby leading; thereby leading to increased perceived *time* while crossing boundary nodes relative to staying within the cluster. Following the findings in chapter 2, this increased perceived time can be modeled through entropy difference between boundary and non-boundary nodes based on learned SR representations. This slowdown is simulated by multiplying average SR entropy of boundary nodes where they transitions cross boundaries and by multiplying the average SR entropy of non-boundary nodes where transitions stay within a cluster. predictions of differences in perceived distance are shown in figure 3.3.

Regardless of whether the distance measure is boosted by increased entropy at boundary nodes, it appears that boundary nodes themselves appear farther from each other relative to nodes within a cluster from that cluster's boundary node. Model predictions at other distances are mixed and dependent on parameters of the SR model. The next experiment tests whether the typical finding of increased perceived separation between cross cluster nodes (relative to within cluster nodes) is replicated in implicitly operationalized event boundary paradigms. A rigorous test of this formulation of the SR model is out of the scope of this dissertation. Furthermore, a lack of increased cross-cluster distance will provide no evidence for or against the current formulation of the SR model's role in representing distances between nodes. However, if such an increased cross-cluster distance is found, such an effect will serve as evidence *against* the current formulation of the SR model's role in estimating temporal distances.

3.4 Discussion

As exemplified above, this modeling approach allows for varying accounts of memory improvement for implicitly operationalized boundary nodes relative to non-boundary nodes. This approach does not allow to differentiate between these accounts. The focus of this chapter is to demonstrate whether implicitly operationalized boundary items are remembered better than non-boundary items similar to findings in explicitly operationalized event boundary literature. The increased memory strength assumption is not

DDM is one model. EBRW is a more direct extension of GCM. Cite cohen.

CHAPTER 4

CATEGORY LEARNING THROUGH TEMPORAL ABSTRACTION

4.1 (Partly polished) Introduction

We naturally categorize items we encounter daily for ease of storage, processing, and decision-making. For example we know instinctively that regardless of shape and form, all lamps form a 'lamp' category based on its function. Several factors determine how we categorize items. Depending on the complexity of rules that determine categories, some categorizations are easier than others [72, 57]. Category variability can modulate how often exemplars are classified into that categories [14].

The order of presentation items in category learning tasks has been shown to be an important factor in how category diagnostic features are learned. In particular, when items are presented in a blocked categorical design, participants seem to learn the similarities between the same category items. On the other hand, when items are presented as an interleaved design, participants seem to focus more on learning the features that differentiate the underlying categories [11]. As a result of order-dependent differing focus on category diagnostic features, interleaved presentations seem to benefit general category learning. In most prior category-learning tasks assessing order of presentation effects, participants are explicitly asked to learn the underlying categories. There appear to be clear differences when participants focus on learning categories based on how exemplars of these categories are presented [39, 40, 86, 84, 10, 11]. In this article, we investigate the effects of order of presentation when category learning is implicit.

One primary focus on category learning through order of presentation is comparing blocked or interleaved exemplar presentations. For example, [39] showed participants paintings made by two different painters. The order of presentation during exposure was modulated to either be blocked (paintings of one artist shown together followed by the second artist) or interleaved (paintings made by both artists were mixed). When presented with new paintings, and asked which of the two studied artists made them, participants who were exposed to the interleaved format were found to be more accurate at guessing the creator. Category learning also improved for interleaved presentation compared to blocked presentation when tested on items where relevant category features were visually occluded [86]. When three-year-old children are tested on the generalization of category-specific features, they appear to benefit from the spaced study of exemplars as compared to a blocked [84]. By modulating the similarity of presented items, interleaved presentation was found to be better than blocked presentation design on generalization performance particularly when learned exemplars were more visual [39, 10].

Interleaved presentation has been theorized to improve in category induction because of context-based variability during encoding [29]. Particularly, for each presented item, an observer will store both the item-specific features along with the context in which the item is encoded. During interleaved presentation, a category diagnostic feature gets encoded under different contexts. Thus, that diagnostic feature will be recalled when tested on novel category items within that context.

Two theories have been proposed to explain this discriminability-based advantage of interleaving. According to the attention attenuation account, when categories are blocked, participants may think that they have learned the relevant category features after viewing a few items and stop paying attention to additional exemplars of the [40]. On the other hand, according to the discrimination account, the interleaved presentation allows participants to directly compare the differences between exemplars of different categories that are presented close to each other thereby highlighting these differences [39]. In a direct test [85] found that when participants were shown pairs of exemplars, each belonging to a different category, the interleaving benefit was magnified compared to when they were presented as single items. The authors posit that showing pairs of exemplars would enable participants to carefully study and infer distinctions between category features and hence improve categorization performance .Furthermore, the authors find evidence against the attention attenuation theory by observing that classification performance did not differ as a function of the position in which the exemplar was presented in a stream.

This benefit of interleaved presentation is shown to be modulated by the 'level' at which categorization occurs. For example, when [48] modulated exposure time to individual exemplars along with order of exposure, they found that interleaved presentation was no longer beneficial under short exposure conditions particularly when participants were asked to make a more abstract, 'super-ordinate level categorization. On the other hand, when exemplars were presented in a blocked format, a lower, 'basic' level categorization was hindered. Thus, category knowledge through order of presentation can be modulated by the level of categorization participants are asked to produce.

It is clear that order of presentation of categories matters during explicit category learning. **However, the effect of such order of presentation has not been investigated when category learning is implicit**. Indeed recent work shows that participants do acquire category knowledge that when presented implicitly instead of being explicitly asked to learn categories. [82] found that assessed on category knowledge, participants appeared to learn category structures without being explicitly instructed to do so. This category knowledge was modulated by the strength of association of the category diagnostic features. [83] later found that when presented with implicit feature-based categories during a cover task, participants were sensitive towards category diagnostic knowledge.

More evidence for incidental category learning comes from auditory cognition. [25] found that participants were sensitive to audio categories learned implicitly as measured by increased reaction times when audio-category-to-response mapping was altered. Incidental category knowledge is further modulated by the sampling category distributions from which exemplars are drawn. [64] show that probabilistic sampling of exemplars leads to weaker category learning compared to deterministic sampling. Incidental category learning can be further enhanced by task-relevant and disrupted by task-irrelevant feature-to-category mappings [65]. Incidental learning may also be disadvantaged compared to supervised intentional learning when categories are non-linearly separable [43].

Thus, while implicit category learning appears to be consistent and dependent on several aspects of the underlying categories, unlike explicit category learning, it is unclear whether implicit category learning enjoys the same advantage when category exemplars are presented in an interleaved vs. a blocked design. Most implicit categorization tasks involve manipulation of features as opposed to manipulation of the temporal order of exposure. In this article, we explore the effects of temporal co-occurrences to manipulate

4.1.0.0.1 Other stuff to incorporate: Unsupervised category learning [6], participants could infer rules based on correlating features without being explicitly asked to categorize during exposure.

Category-related items were recognized better when presented close to each other then when category-unrelated items were [49]. In [50], people unsupervised sorted stimuli by a single dimension, ignoring the family resemblance structure.

SUSTAIN [44] seems to explain all these unsupervised category learning phenomena. ALCOVE [41] provides for error based diagnostic feature attention learning in GCM [56, 55].

4.1.1 Participants

4.1.2 Materials

CHAPTER 5

DISCUSSION

BIBLIOGRAPHY

[1] Aslin, Richard N, and Newport, Elissa L. Statistical learning: From acquiring specific items to forming general rules. *Current directions in psychological science 21*, 3 (2012), 170–176.

[2] Baldassano, Christopher, Chen, Janice, Zadbood, Asieh, Pillow, Jonathan W, Hasson, Uri, and Norman, Kenneth A. Discovering event structure in continuous narrative perception and memory. *Neuron 95*, 3 (2017), 709–721.

[3] Baldwin, Dare, Andersson, Annika, Saffran, Jenny, and Meyer, Meredith. Segmenting dynamic human action via statistical structure. *Cognition 106*, 3 (2008), 1382–1407.

[4] Behrens, Timothy EJ, Woolrich, Mark W, Walton, Mark E, and Rushworth, Matthew FS. Learning the value of information in an uncertain world. *Nature neuroscience 10*, 9 (2007), 1214–1221.

[5] Bera, Krishn, Shukla, Anuj, and Bapi, Raju S. Motor chunking in internally guided sequencing. *Brain Sciences 11*, 3 (2021), 292.

[6] Billman, Dorrit, and Knutson, James. Unsupervised concept learning and value systematicitiy: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, and Cognition 22*, 2 (1996), 458.

[7] Botvinick, Matthew Michael. Hierarchical reinforcement learning and decision making. *Current opinion in neurobiology 22*, 6 (2012), 956–962.

[8] Brady, Timothy F, and Oliva, Aude. Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological science 19*, 7 (2008), 678–685.

[9] Brunec, Iva K, Moscovitch, Morris, and Barense, Morgan D. Boundaries shape cognitive representations of spaces and events. *Trends in cognitive sciences 22*, 7 (2018), 637–650.

[10] Carvalho, Paulo F, and Goldstone, Robert L. Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & cognition 42* (2014), 481–495.

[11] Carvalho, Paulo F, and Goldstone, Robert L. The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition 43*, 11 (2017), 1699.

[12] Cleeremans, Axel, and McClelland, James L. Learning the structure of event sequences. *Journal of Experimental Psychology: General 120*, 3 (1991), 235.

[13] Clewett, David, DuBrow, Sarah, and Davachi, Lila. Transcending time in the brain: How event memories are constructed from experience. *Hippocampus 29*, 3 (2019), 162–183.

[14] Cohen, Andrew L, Nosofsky, Robert M, and Zaki, Safa R. Category variability, exemplar similarity, and perceptual classification. *Memory & Cognition 29*, 8 (2001), 1165–1175.

[15] Davis, Tyler, Love, Bradley C, and Preston, Alison R. Striatal and hippocampal entropy and recognition signals in category learning: simultaneous processes revealed by model-based fmri. *Journal of Experimental Psychology: Learning, Memory, and Cognition 38*, 4 (2012), 821.

[16] Dayan, Peter. Improving generalization for temporal difference learning: The successor representation. *Neural computation 5*, 4 (1993), 613–624.

[17] Dehaene, Stanislas, Meyniel, Florent, Wacongne, Catherine, Wang, Liping, and Pallier, Christophe. The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron 88*, 1 (2015), 2–19.

[18] Dezfouli, Amir, and Balleine, Bernard W. Habits, action sequences and reinforcement learning. *European Journal of Neuroscience 35*, 7 (2012), 1036–1051.

[19] Dezfouli, Amir, Lingawi, Nura W, and Balleine, Bernard W. Habits as action sequences: hierarchical action control and changes in outcome value. *Philosophical Transactions of the Royal Society B: Biological Sciences 369*, 1655 (2014), 20130482.

[20] Dolan, Ray J, and Dayan, Peter. Goals and habits in the brain. *Neuron 80*, 2 (2013), 312–325.

[21] DuBrow, Sarah, and Davachi, Lila. The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology: General 142*, 4 (2013), 1277.

[22] Ezzyat, Youssef, and Davachi, Lila. What constitutes an episode in episodic memory? *Psychological science 22*, 2 (2011), 243–252.

[23] Ezzyat, Youssef, and Davachi, Lila. Similarity breeds proximity: pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron 81*, 5 (2014), 1179–1189.

[24] Fiser, József, and Aslin, Richard N. Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition 28*, 3 (2002), 458.

[25] Gabay, Yafit, Dick, Frederic K, Zevin, Jason D, and Holt, Lori L. Incidental auditory category learning. *Journal of Experimental Psychology: Human Perception and Performance 41*, 4 (2015), 1124.

[26] Gershman, Samuel J. The successor representation: its computational logic and neural substrates. *Journal of Neuroscience 38*, 33 (2018), 7193–7200.

[27] Gershman, Samuel J, Moore, Christopher D, Todd, Michael T, Norman, Kenneth A, and Sederberg, Per B. The successor representation and temporal context. *Neural Computation 24*, 6 (2012), 1553–1568.

[28] Gershman, Samuel J, and Niv, Yael. Learning latent structure: carving nature at its joints. *Current opinion in neurobiology 20*, 2 (2010), 251–256.

[29] Glenberg, Arthur M. Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition 7*, 2 (1979), 95–112.

[30] Hahamy, Avital, Dubossarsky, Haim, and Behrens, Timothy EJ. The human brain reactivates context-specific past information at event boundaries of naturalistic experiences. *Nature neuroscience 26*, 6 (2023), 1080–1089.

[31] Heusser, Andrew C, Ezzyat, Youssef, Shiff, Ilana, and Davachi, Lila. Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology: Learning, Memory, and Cognition 44*, 7 (2018), 1075.

[32] Horner, Aidan J, Bisby, James A, Wang, Aijing, Bogus, Katrina, and Burgess, Neil. The role of spatial boundaries in shaping long-term event representations. *Cognition 154* (2016), 151–164.

[33] Howard, Marc W, Fotedar, Mrigankka S, Datey, Aditya V, and Hasselmo, Michael E. The temporal context model in spatial navigation and relational learning: toward a common explanation of medial temporal lobe function across domains. *Psychological review 112*, 1 (2005), 75.

[34] Kahn, Ari E, Karuza, Elisabeth A, Vettel, Jean M, and Bassett, Danielle S. Network constraints on learnability of probabilistic motor sequences. *Nature human behaviour 2*, 12 (2018), 936–947.

[35] Karuza, Elisabeth A. The value of statistical learning to cognitive network science. *Topics in Cognitive Science 14*, 1 (2022), 78–92.

[36] Karuza, Elisabeth A, Kahn, Ari E, and Bassett, Danielle S. Human sensitivity to community structure is robust to topological variation. *Complexity 2019*, 1 (2019), 8379321.

[37] Karuza, Elisabeth A, Kahn, Ari E, Thompson-Schill, Sharon L, and Bassett, Danielle S. Process reveals structure: How a network is traversed mediates expectations about its architecture. *Scientific reports 7*, 1 (2017), 12733.

[38] Knowlton, Barbara J, Ramus, Seth J, and Squire, Larry R. Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological science 3*, 3 (1992), 172–179.

[39] Kornell, Nate, and Bjork, Robert A. Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological science 19*, 6 (2008), 585–592.

[40] Kornell, Nate, Castel, Alan D, Eich, Teal S, and Bjork, Robert A. Spacing as the friend of both memory and induction in young and older adults. *Psychology and aging 25*, 2 (2010), 498.

[41] Kruschke, John K. Alcove: An exemplar-based connectionist model of category learning. In *Connectionist Psychology*. Psychology Press, 2020, pp. 107–138.

[42] Lositsky, Olga, Chen, Janice, Toker, Daniel, Honey, Christopher J, Shvartsman, Michael, Poppenk, Jordan L, Hasson, Uri, and Norman, Kenneth A. Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *elife 5* (2016), e16070.

[43] Love, Bradley C. Comparing supervised and unsupervised category learning. *Psychonomic bulletin & review 9*, 4 (2002), 829–835.

[44] Love, Bradley C, Medin, Douglas L, and Gureckis, Todd M. Sustain: a network model of category learning. *Psychological review 111*, 2 (2004), 309.

[45] Lynn, Christopher W, and Bassett, Danielle S. How humans learn and represent networks. *Proceedings of the National Academy of Sciences 117*, 47 (2020), 29407–29415.

[46] Lynn, Christopher W, Kahn, Ari E, Nyema, Nathaniel, and Bassett, Danielle S. Abstract representations of events arise from mental errors in learning and memory. *Nature communications 11*, 1 (2020), 2313.

[47] Lynn, Christopher W, Papadopoulos, Lia, Kahn, Ari E, and Bassett, Danielle S. Human information processing in complex networks. *Nature Physics 16*, 9 (2020), 965–973.

[48] Mack, Michael L, and Palmeri, Thomas J. The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology: General 144*, 3 (2015), 551.

[49] Medin, Douglas L, and Bettger, Jeffrey G. Presentation order and recognition of categorically related examples. *Psychonomic bulletin & review 1*, 2 (1994), 250–254.

[50] Medin, Douglas L, Wattenmaker, William D, and Hampson, Sarah E. Family resemblance, conceptual cohesiveness, and category construction. *Cognitive psychology 19*, 2 (1987), 242–279.

[51] Michelmann, Sebastian, Hasson, Uri, and Norman, Kenneth A. Evidence that event boundaries are access points for memory retrieval. *Psychological Science 34*, 3 (2023), 326–344.

[52] Momennejad, Ida. Learning structures: predictive representations, replay, and generalization. *Current Opinion in Behavioral Sciences 32* (2020), 155–166.

[53] Momennejad, Ida, Russek, Evan M, Cheong, Jin H, Botvinick, Matthew M, Daw, Nathaniel Douglass, and Gershman, Samuel J. The successor representation in human reinforcement learning. *Nature human behaviour 1*, 9 (2017), 680–692.

[54] Nissen, Mary Jo, and Bullemer, Peter. Attentional requirements of learning: Evidence from performance measures. *Cognitive psychology 19*, 1 (1987), 1–32.

[55] Nosofsky, Robert M. Attention, similarity, and the identification–categorization relationship. *Journal of experimental psychology: General 115*, 1 (1986), 39.

[56] Nosofsky, Robert M. The generalized context model: An exemplar model of classification. *Formal approaches in categorization* (2011), 18–39.

[57] Nosofsky, Robert M, Gluck, Mark A, Palmeri, Thomas J, McKinley, Stephen C, and Glauthier, Paul. Comparing modes of rule-based classification learning: A replication and extension of shepard, hovland, and jenkins (1961). *Memory & cognition 22*, 3 (1994), 352–369.

[58] Nosofsky, Robert M, Little, Daniel R, Donkin, Christopher, and Fific, Mario. Short-term memory scanning viewed as exemplar-based categorization. *Psychological review 118*, 2 (2011), 280.

[59] Ostlund, Sean B, Winterbauer, Neil E, and Balleine, Bernard W. Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *Journal of Neuroscience 29*, 25 (2009), 8280–8287.

[60] Radvansky, Gabriel A, and Zacks, Jeffrey M. Event boundaries in memory and cognition. *Current opinion in behavioral sciences 17* (2017), 133–140.

[61] Ratcliff, Roger, Scharre, Douglas W, and McKoon, Gail. Discriminating memory disordered patients from controls using diffusion model parameters from recognition memory. *Journal of Experimental Psychology: General 151*, 6 (2022), 1377.

[62] Ratcliff, Roger, and Starns, Jeffrey J. Modeling confidence and response time in recognition memory. *Psychological review 116*, 1 (2009), 59.

[63] Ratcliff, Roger, Thapar, Anjali, and McKoon, Gail. A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language 50*, 4 (2004), 408–424.

[64] Roark, Casey L, and Holt, Lori L. Task and distribution sampling affect auditory category learning. *Attention, Perception, & Psychophysics 80* (2018), 1804–1822.

[65] Roark, Casey L, Lehet, Matthew I, Dick, Frederic, and Holt, Lori L. The representational glue for incidental category learning is alignment with task-relevant behavior. *Journal of Experimental Psychology: Learning, Memory, and Cognition 48*, 6 (2022), 769.

[66] Romberg, Alexa R, and Saffran, Jenny R. Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science 1*, 6 (2010), 906–914.

[67] Rouhani, Nina, Norman, Kenneth A, and Niv, Yael. Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition 44*, 9 (2018), 1430.

[68] Rouhani, Nina, Norman, Kenneth A, Niv, Yael, and Bornstein, Aaron M. Reward prediction errors create event boundaries in memory. *Cognition 203* (2020), 104269.

[69] Russek, Evan M, Momennejad, Ida, Botvinick, Matthew M, Gershman, Samuel J, and Daw, Nathaniel D. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS computational biology 13*, 9 (2017), e1005768.

[70] Savalia, Tejas, Shukla, Anuj, and Bapi, Raju S. A unified theoretical framework for cognitive sequencing. *Frontiers in psychology 7* (2016), 1821.

[71] Schapiro, Anna C, Rogers, Timothy T, Cordova, Natalia I, Turk-Browne, Nicholas B, and Botvinick, Matthew M. Neural representations of events arise from temporal community structure. *Nature neuroscience 16*, 4 (2013), 486–492.

[72] Shepard, Roger N, Hovland, Carl I, and Jenkins, Herbert M. Learning and memorization of classifications. *Psychological monographs: General and applied 75*, 13 (1961), 1.

[73] Shin, Yeon Soon, and DuBrow, Sarah. Structuring memory through inference-based event segmentation. *Topics in Cognitive Science 13*, 1 (2021), 106–127.

[74] Smith, Kyle S, and Graybiel, Ann M. Habit formation. *Dialogues in clinical neuroscience 18*, 1 (2016), 33–43.

[75] Sols, Ignasi, DuBrow, Sarah, Davachi, Lila, and Fuentemilla, Lluís. Event boundaries trigger rapid memory reinstatement of the prior events to promote their representation in long-term memory. *Current Biology 27*, 22 (2017), 3499–3504.

[76] Stachenfeld, Kimberly L, Botvinick, Matthew M, and Gershman, Samuel J. The hippocampus as a predictive map. *Nature neuroscience 20*, 11 (2017), 1643–1653.

[77] Starns, Jeffrey J. Using response time modeling to distinguish memory and decision processes in recognition and source tasks. *Memory & cognition 42* (2014), 1357–1372.

[78] Starns, Jeffrey J, and Ratcliff, Roger. Validating the unequal-variance assumption in recognition memory using response time distributions instead of roc functions: A diffusion model analysis. *Journal of memory and language 70* (2014), 36–52.

[79] Swallow, Khena M, Zacks, Jeffrey M, and Abrams, Richard A. Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General 138*, 2 (2009), 236.

[80] Tremblay, Pierre-Luc, Bedard, Marc-Andre, Langlois, Dominic, Blanchet, Pierre J, Lemay, Martin, and Parent, Maxime. Movement chunking during sequence learning is a dopamine-dependant process: a study conducted in parkinson's disease. *Experimental brain research 205* (2010), 375–385.

[81] Turk-Browne, Nicholas B, Isola, Phillip J, Scholl, Brian J, and Treat, Teresa A. Multidimensional visual statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition 34*, 2 (2008), 399.

[82] Unger, Layla, and Sloutsky, Vladimir M. Ready to learn: Incidental exposure fosters category learning. *Psychological Science 33*, 6 (2022), 999–1019.

[83] Unger, Layla, Weichart, Emily R, Reardon, Noah, and Sloutsky, Vladimir. Without even trying: How incidental exposure shapes category learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (2023), vol. 45.

[84] Vlach, Haley A, Sandhofer, Catherine M, and Kornell, Nate. The spacing effect in children's memory and category induction. *Cognition 109*, 1 (2008), 163–167.

[85] Wahlheim, Christopher N, Dunlosky, John, and Jacoby, Larry L. Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & cognition 39* (2011), 750–763.

[86] Whitehead, Peter S, Zamary, Amanda, and Marsh, Elizabeth J. Transfer of category learning to impoverished contexts. *Psychonomic Bulletin & Review* (2021), 1–10.

[87] Zacks, Jeffrey M. Event perception and memory. *Annual review of psychology 71*, 1 (2020), 165–191.

[88] Zacks, Jeffrey M, Speer, Nicole K, Swallow, Khena M, Braver, Todd S, and Reynolds, Jeremy R. Event perception: a mind-brain perspective. *Psychological bulletin 133*, 2 (2007), 273.

[89] Zacks, Jeffrey M, and Swallow, Khena M. Event segmentation. *Current directions in psychological science 16*, 2 (2007), 80–84.