Measuring Components of Order Memory

A thesis submitted by

Giulia Isabella Pintea

in partial fulfillment of the requirements for the degree of

Master of Science

in

Psychology

Tufts University

Graduate School of Arts and Sciences

May 2022

Advisor: Richard Chechile, PhD

Abstract

The present research is based on the multiple-attribute memory framework. According to this framework, there are different aspects to the memory for an event. These aspects include the content of the stimulus as well as temporal and ordinal information pertaining to the stimulus. The key questions in this research follow from this framework: Are content information and order information unitized? (Hinrichs, 1970) Or are they attributes with different learning and forgetting properties? (Underwood, 1977). Furthermore, how can these two aspects of event memory be measured? And what model can account for the learning of order?

Using a novel task that probes order knowledge of a triplet of items, we were able to measure four states of order knowledge: (1) knowledge of all three items in the triplet, (2) knowledge of two of the order relationships among the three items, (3) knowledge of a single order relationship, and (4) no order knowledge. These knowledge states are measured by the following probabilities: θ_{3C}, θ_{2C}, θ_{1C}, and θ_N. The four experiments in this thesis explore how these parameters vary as a function of training trials, spacing of the items, and the relativeness of the items in the triplet. We also employed a separate model previously developed by Chechile and Soraci (1999) to measure the probability of item storage.

This study is organized as follows. Experiment 1 was our pilot study, where we tested participants on the order of stimuli following two training trials. In Experiment 2, we increased the number of training trials from two to four. In Experiment 3, we included item-storage questions to compare performance on item memory vs. order memory. Finally, in Experiment 4 we increased the number of training trials from four to eight. The four experiments showed that learning of order is more than simply learning the items in a triplet. Order learning of the elements of a triplet improves when the items are related, and there are marked individual differences for this type of learning. Furthermore, the rate of improvement after the initial trial slows down with further training, which we were able to describe using Weibull modeling.

Dedication

I dedicate this dissertation to a few beloved people who have meant and continue to mean so much to me. First and foremost, to Rae'Niqua, who has supported me and taught me the importance of trusting the process. Secondly, my undergraduate professor, Dr. Beers, who has always encouraged me to pursue my dreams. And thirdly, I dedicate this thesis to my partner, David, who has been a constant source of encouragement during the challenges of graduate school and life. I am truly thankful for having you in my life.

Acknowledgments

I would like to express my deepest thanks to the people who have made this work possible. Firstly, I would like to thank my advisor, Dr. Richard Chechile, for providing guidance and feedback throughout this project. Secondly, I would like to thank the Memory Lab RAs that contributed to the data collection process: Claire Foley, Dominique Landinez, Amy Ly, Gillian Roeca, and Tiffany Wu. Thirdly, I would like to thank Dr. Dan Barch, Dr. Heather Urry, and Dr. Ayanna Thomas for supporting me over the past three years. Finally, my biggest thanks to my partner, David – I couldn't have done this without you.

List of Tables

List of Figures

List of Copyrighted Materials Used

No copyrighted materials were used in this dissertation.

List of Abbreviations

AFC .... Alternative Forced-Choice
AR .... Adjacent-Related
AU .... Adjacent-Unrelated
BF .... Bayes Factor
EO .... Event Order
FR .... Far-Related
FU .... Far-Unrelated
IS .... Item Storage
MLE .... Maximum Likelihood Estimator
MPT .... Multinomial Processing Tree
NR .... Near-Related
NU .... Near-Unrelated
PPM .... Population Parameter Mapping

Contents

Chapter 1

Modeling How to Measure Order Memory

1.0.1 Introduction

Much of the research conducted in the memory area has focused on the content of memory. Content, however, is just one of the multiple attributes of memory. The context in which memories are stored plays a role in later recall, and context itself can have different features.

For example, memory context can refer to time and order. Benton J. Underwood (1977) investigated attributes of memory outside of content, such as the event frequency associated with that memory, the duration of exposure to that event, as well as the role of interference in temporal coding. Despite Underwood's research, the focus in memory research has nonetheless been on the content (or ”what”) of an episode rather than on the other attributes, such as the ”when,” ”where,” and ”who.”

When a new memory is stored, there are many aspects of said memory that are processed for storage. Such aspects include location, emotion, the primary subject(s) of the memory, and more. In our research we focus on the ordinal information that is stored as a part of learning. Remembering the order a series of memories occurred in is crucial for our survival and day-to-day functioning. In the late twentieth century, several experimental procedures in psychological research were implemented to explore temporal and ordinal knowledge. Such procedures include: subjective judgement of elapsed time (Hicks et al., 1976), subjective judgment of the relative duration between two remembered intervals (Block, 1986), subjective estimation of the lag between item repetitions (Hinrichs and Buschke, 1968), judgment of the position of recalled items (Michon and Jackson, 1984), dating of long-term world events (Underwood, 1977), selection of the more recent of two provided items (Fozard, 1970), and rearranging items provided in a random order to match the original order (Shimamura et al., 1990).

The results that were obtained from these experimental procedures led some researchers to believe that temporal information was encoded in memory. For example, Underwood (1977) conducted a study where item knowledge for eight world events was excellent; however, only two of the 108 participants who participated in the study could correctly rank-order all the events. Simultaneously, the overall rank-order of the participants was correlated with the correct order of the world events. Based on these findings, Underwood (1983, 1977) argued that a temporal code could be developed as a separate attribute of memory throughout the learning process. In other words, the learning of order was additional to the learning of items themselves. Estes (1985), argued that the strength model introduced by Hinrichs (1970) could potentially account for data such as judgments of the relative recency of events. In strength theory, each item is presumed to have an associated strength that decays over time, and it is assumed that items with stronger memory traces are judged as being the most recent ones. Although Estes (1985) preferred a perturbation model instead (Estes, 1985, 1972; Lee, 1992; Lee and Estes, 1981), he acknowledged that a strength model would make a creditable account, despite the fact that it did not account for encoding either temporal or ordinal information. Based on this, we can argue that none of the existing behavioral procedures in the twentieth century have provided any evidence that people were encoding temporal or ordinal information as part of the learning process.

We can take a look at neuroscientific research studies to obtain more insights on the storage of temporal information in the brain. We know that the hippocampus plays an important role in memory, and thanks to studies that focused on amnesic patients, we also know that the hippocampus is responsible for binding different components of memories that are produced during the encoding of an event (Chechile, 2018). Furthermore, hippocampal cells have been discovered to be in charge of information related to timing signals, which help subjects with their performance on memory tasks concerned with the temporal order of visual stimuli (Naya and Suzuki, 2011). Overall, these cells are sensitive to both sequential and ordinal information (Eichenbaum, 2014; MacDonald et al., 2011; Pastalkova et al., 2008). Outside of the hippocampus, the frontal lobes have also been shown to play a big role in the processing of ordinal knowledge (Kesner et al., 1994; Mangels, 1997; Milner et al., 1991; Shimamura et al., 1990). The use of event-related fMRI scans on humans has also helped researchers gather evidence that both the temporal lobe and the prefrontal cortex are active when subjects undergo tasks that require the processing of temporal information (Jenkins and Ranganath, 2010). In addition, brain rhythms have been identified, which were shown to be active when animals were engaged in memory tasks (O'Keefe and Recce, 1993). Furthermore, Wearden and Towse (1994); Wearden and Bray (2001) showed that the various, co-occurring brain rhythms form a possible basis for obtaining a temporal encoding for events. This possibility was endorsed in the OSCAR computational model for serial order (Brown et al., 2000). Thus, we have a body of neuroscientific research studies that provides us with evidence of a biological basis for the integration of different components of memory, particularly those concerned with temporal and ordinal information.

From a theoretical angle, there are some researchers in the twenty-first century who have argued that episodic memory is organized along a temporal dimension (Brown et al., 2007, 2000; Chechile, 2018; Howard et al., 2015, 2014; Howard and Eichenbaum, 2015; Ward et al., 2013). However, Howard et al. (2015) stressed that the temporal flow of memory is not at all uniform; it can have jumps and breaks due to the recall of an item triggering a prior time, which as a result can trigger other content information. Moreover, it is possible to scan memory by using time as the primary searching tool.

Although the temporal coding of information is a fair proposition, it is not yet established that time is a key feature of memory organization and storage. To date, there is no consensus in psychology over the process by which people and animals develop a mental model of events that reflects the order in which said events took place. Some researchers (Friedman, 1993) have argued that experimental data can be accounted for by hypotheses that do not involve temporal code. However, despite this lack of consensus, we can make progress in our understanding of order memory by working towards improved psychometric methods. The existing dependent variables for studying the order of events are limited because they do not account for guessing factors and they do not account for several stages or degrees of order knowledge. There are many other examples in cognitive psychology where the dependent variables involve an entangled combination of latent processes that cannot be directly studied individually. However, with model-based measurement, these latent processes can be studied, and hence, measured. See Chechile (2018) for a more general review of how Multinomial Processing Tree (MPT) models can be used to improve the toolkit of memory researchers by extracting underlying processes from the available behavioral measures. MPT models also have desirable statistical properties that provide for a distribution-free analysis of differences between groups or conditions. Useful even for a single experimental condition, MPT models provide for an informative measure of the separate latent processes. The current paper is focused on the development of such a memory measurement tool, which we refer to as the Event Order (EO) model. The EO model is used, along with other modeling tools, to provide evidence that order knowledge is a separate memory attribute or feature of the memory representation from the other stimulus aspects of an event.

The remainder to the paper is roughly divided into: (1) the development and discussion of the event-order (EO) model; (2) the description of four experiments that use the EO model; (3) a discussion of the Weibull model for learning order information; and (4) an overall discussion of the key findings.

1.0.2 Modeling Order

Memory Task

The eventual multinomial model for measuring the order of events is based on a specific experimental task. The task consists of a learning phase that is followed by a testing phase. The stimuli in the learning phase are presented one at a time. Our stimuli of choice are words, but other stimuli are also possible with this experimental task. In the testing phase of a trial, the participants are presented a series of forced-choice tests about the order among three selected items from the original set of stimuli [1].

The multiple-choice questions of the testing phase allowed the participants to have three attempts to get the correct answer. The testing for a set of three stimuli (denoted here as ABC) is conducted by a series of alternative forced-choice (AFC) questions that is illustrated in Figure 1.1. For the first attempt, the participants were presented with a 6-AFC task, as they had six options to pick the correct answer from. If they answered correctly on the first attempt, the trial ended and they moved on to the next question. If they answered incorrectly, their incorrect answer was removed and they tried again on a 5-AFC. If they answer incorrectly on this second attempt, their incorrect guess is removed, such that they can try again on a 4-AFC. On this third (and last) attempt, the trial ends regardless of whether their answer is correct or incorrect, and they move on to the next question. Thus, a trial test results in one of four categorical outcomes (i.e., a cell in a four-cell multinomial). After a number of triplets of stimuli are tested in a similar fashion, the testing phase for that trial is concluded.

This task is similar to the one used by Chechile and Soraci (1999), who used AFC tasks to obtain model-based measurements of storage and retrieval. Chechile et al. (2012) also used a similar contingent series of AFC tests to obtain separate estimates of explicit and implicit memory, which called the implicit-explicit separation (IES) model. The data used to construct this model came from a 4-AFC task (Chechile,

[1] If there were only two items selected then the subsequent multinomial model of order would not have enough information. Moreover, if four stimuli were probed for testing then the time to test the participant would be too costly.

Visual Representation of the AFC Task

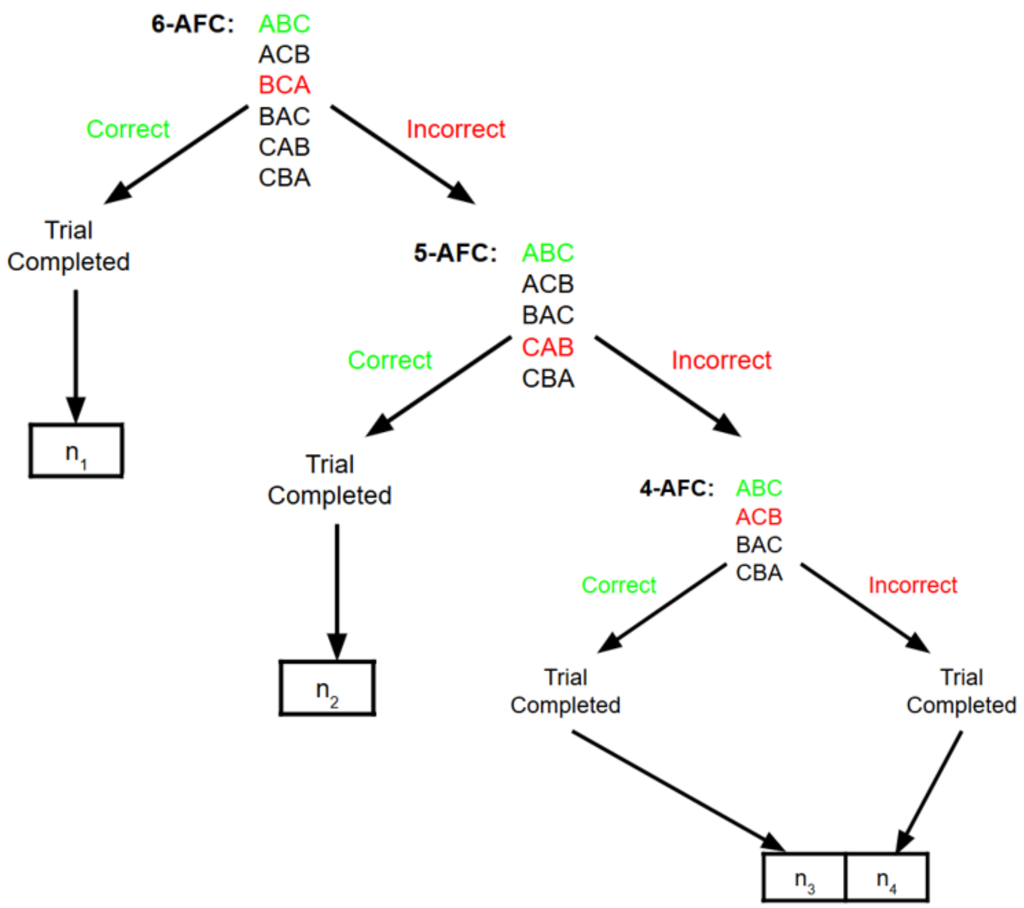

Figure 1.1: Visual illustration of the AFC task in the testing phase for the memory task of the study. The correct answer (ABC) is shown in green, and the (hypothetical) incorrect attempts are shown in red. The values of n correspond to the frequencies we later use for our model.

2018) that follows a structure similar to the one pictured in Figure 1.1.

Multinomial Processing Tree (MPT) Model

Multinomial Processing Tree (MPT) models are used to investigate entangled cognitive processes that cannot be directly measured. The model is based on the frequencies in the four categories delineated in Figure 1.1. The pertinent of a four-cell multinomial are the following:

n_1 = frequency of correct attempts on the 6-AFC

n_2 = frequency of correct attempts on the 5-AFC

n_3 = frequency of correct attempts on the 4-AFC

n_4 = frequency of incorrect attempts on the 4-AFC

From these frequencies, we obtain the multinomial probabilities of our model. We denote the population probabilities for the corresponding for cells as ϕ_i, where $i = 1, ..., 4$, where the sum of all four probabilities is 1.

ϕ_1 = probability of answering correctly on the 6-AFC

ϕ_2 = probability of answering correctly on the 5-AFC

ϕ_3 = probability of answering correctly on the 4-AFC

ϕ_4 = probability of answering incorrectly on the 4-AFC

The goal of our MPT model is to estimate the probabilities for the four different levels of knowledge of the order of the stimuli:

$\Theta = (\theta_{3C}, \theta_{2C}, \theta_{1C}, \theta_N)$

θ_{3C} = knowledge of the order of all 3 stimuli

θ_{2C} = knowledge of the order of 2 of 3 stimuli

θ_{1C} = knowledge of 1 of the 3 order relationships

θ_N = no knowledge of the order of stimuli

Where $\theta_{3C} + \theta_{2C} + \theta_{1C} + \theta_N = 1$.

See Figure 1.2 for the tree representation for the EO model. This tree representation shows the probabilities of four categorical outcomes (cell 1, cell 2, cell 3, and cell 4) given a particular memory state (θ_{3C}, θ_{2C}, θ_{1C}, and θ_N).

If the subject is assumed to have knowledge of the order of all three items (θ_{3C}), then he or she will answer the 6-AFC task correctly, thus resulting in the cell 1 outcome. If the subject is assumed to have knowledge of the order of only two of the three items (θ_{2C}), then there are two possible outcomes. The participant will either answer the 6-AFC correctly (cell 1), which has a probability of 1/2. Or, alternatively, with a probability of 1/2, the participant will be incorrect on the 6-AFC task and go on to be correct on the 5-AFC task (i.e., a cell 2 response). If the subject is assumed to have knowledge of the order of only one of the three items (θ_{1C}), then there are three possible outcomes. The participant could answer the 6-AFC task correctly (cell 1), with a 1/3 probability. However, if he or she answers incorrectly (2/3 probability), the participant goes on to a second attempt. Now, with a 5-AFC task, the participant has a 1/2 chance of answering correctly (cell 2), or a 1/2 chance of answering incorrectly and subsequently answer correctly on the 4-AFC task (cell 3).

If the subject is assumed to have no knowledge of order whatsoever (θ_N) there are four possible outcomes. The participant has a 1/6 chance of answering correctly on the 6-AFC task task (cell 1), and a 5/6 chance of answering incorrectly and moving on to the 5-AFC task. On the 5-AFC task, the participant now has a 1/5 chance of answering correctly (cell 2) and a 4/5 chance of answering incorrectly and moving on to the 4-AFC task. On the 4-AFC task, the participant has a 1/4 chance of answering correctly (cell 3) and a 3/4 chance of answering incorrectly (cell 4).

Illustration of the EO Measurement Model

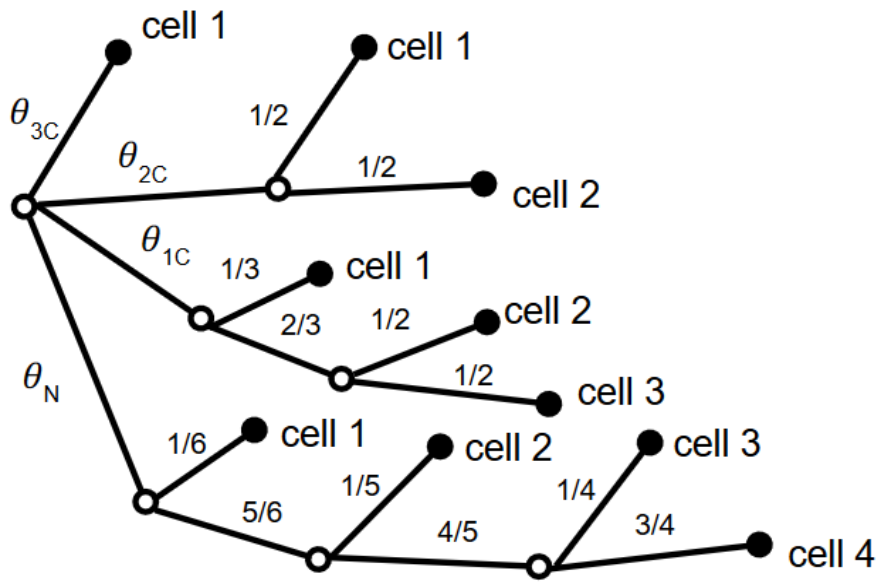

Figure 1.2: An illustration of the structure of the event-order (EO) measurement model for the contingent foced-choice test. The outcome cells correspond to the four categorical outcomes that have frequencies n_1, n_2, n_3, and n_4.

From the processing tree in Figure 1.2, the probabilities for the four response cells can be expressed in terms of the probabilities for the four knowledge states. They are the following:

$$\phi_1 = \frac{\theta_N}{6} + \frac{\theta_{1c}}{3} + \frac{\theta_{2c}}{2} + \theta_{3c} \quad (1.1)$$

$$\phi_2 = \frac{\theta_N}{6} + \frac{\theta_{1c}}{3} + \frac{\theta_{2c}}{2} \quad (1.2)$$

$$\phi_3 = \frac{\theta_N}{6} + \frac{\theta_{1c}}{3} \quad (1.3)$$

$$\phi_4 = \frac{\theta_N}{2} \quad (1.4)$$

With this information, we can set up the likelihood function of our data in the following way:

$$\mathscr{L} = K(\frac{\theta_N}{2})^{n_4}(\frac{\theta_N}{6}+\frac{\theta_{1C}}{3})^{n_3}(\frac{\theta_N}{6}+\frac{\theta_{1C}}{3}+\frac{\theta_{2C}}{2})^{n_2}(1-\frac{\theta_N}{2}-\frac{\theta_N}{6}-\frac{\theta_{1C}}{3}-\frac{\theta_N}{6}-\frac{\theta_{1C}}{3}-\frac{\theta_{2C}}{2})^{n_1}$$

The multinomial likelihood for a four-cell is a generalization of the binomial function. Likelihood is the probability of the data $D = \{n_1, n_2, n_3, n_4\}$ given the stipulated values for the theta parameters, where the sum of these parameters is equal to 1. Finally, in equation above, K is a constant.

Population Parameter Mapping (PPM)

The Population-Parameter Mapping (PPM) method is an approach to parameter estimation that was invented by Chechile in 1998 (Chechile, 1998, 2004, 2010b). The PPM estimation method is a Bayesian Monte Carlo sampling procedure that involves sampling vectors from the posterior distribution for the set of ϕ_i, $\{i = 1, \cdots, 4\}$ and mapping these vectors to a corresponding vector of EO model parameters, provided that the set of ϕ_i are consistent with the EO model. Each sample from the ϕ-space is within a space such that the sum of the ϕ values is 1. In addition, each sample is independent of the EO model. The mapping equations used for this approach are the following:

$$\theta_N = 2\phi_1\,, \quad (1.5)$$

$$\theta_{1C} = 3(\phi_3 - \phi_4)\,, \quad (1.6)$$

$$\theta_{2C} = 2(\phi_2 - \phi_3)\,, \quad (1.7)$$

$$\theta_{3C} = \phi_1 - \phi_2\,, \quad (1.8)$$

Provided that the following four constraints are satisfied for a coherent mapping:

$$\phi_1 \geq \phi_2\,, \quad (1.9)$$

$$\phi_2 \geq \phi_3\,, \quad (1.10)$$

$$\phi_3 \geq \frac{\phi_4}{3}\,, \quad (1.11)$$

$$\phi_4 \leq \frac{1}{2}\,. \quad (1.12)$$

If equation (1.9) is violated, then that would imply the incoherent result of $\theta_{3c} < 0$. If equation (1.10) is violated, then would imply the incoherent result of $\theta_{2c} < 0$. If the equation (1.11) is violated, then that would imply that $\theta_{1c} < 0$. Finally if equation (1.12) is violated, then it would imply that $\theta_N > 1$. The proportion of all the ϕ-space vectors sample that are coherent is a measure of model coherence, which is denoted as $P(coh)$. $P(coh)$ is an estimate of the proportion of the posterior distribution of ϕ-space that is consistent with the EO model. The coherent Monte Carlo samples are used for all subsequent point and interval estimates as well as for statistical tests among experimental conditions. The PPM has been previously shown to be more precise than the maximum likelihood estimator for small sample sizes (Chechile, 2010b).

Maximum Likelihood Estimator (MLE)

The Maximum Likelihood Estimator (MLE) is the joint set of model parameters that maximizes the likelihood function. The MLE is found by solving three simultaneous equations that set the partial derivative of the likelihood function with respect to one of the three independent model parameters to zero.

For the MLE approach, it is necessary to rewrite the likelihood function to be a function of only three of the four θ parameters. This is because the sum of all four parameters equals 1. Hence we set: $\theta_{3C} = 1 - \theta_N - \theta_{1C} - \theta_{2C}$. We then find the maximum of the likelihood as a function of θ_N, θ_{1C}, and θ_{2C}.

Let us consider the case where $n_i > 0$ for $i = \{1, \cdots, 4\}$, the MLE values for the EO parameters are obtained by solving a set of three linear equations that result from the constraints that (1) $\frac{\partial \mathscr{L}}{\partial \theta_{2C}} = 0$, (2) $\frac{\partial \mathscr{L}}{\partial \theta_{1C}} = 0$, and (3) $\frac{\partial \mathscr{L}}{\partial \theta_N} = 0$. These three constraints can be expressed in the following single matrix equation:

$$\mathbf{M} \cdot \begin{pmatrix} \theta_N \\ \theta_{1C} \\ \theta_{2C} \end{pmatrix} = \begin{pmatrix} n_2 \\ 0 \\ 0 \end{pmatrix}, \tag{1.13}$$

Where $\mathbf{M}$ is a 3×3 matrix with elements that are:

$$\begin{aligned}
m_{11} &= \frac{n_1 + 5n_2}{6}, \\
m_{12} &= \frac{n_1 + 2n_2}{3}, \\
m_{13} &= \frac{n_1 + n_2}{2}, \\
m_{21} &= \frac{n_3 - n_2}{6}, \\
m_{22} &= \frac{n_3 - n_2}{3}, \\
m_{23} &= \frac{n_3}{2}, \\
m_{31} &= \frac{n_4 - 3n_3}{6}, \\
m_{32} &= \frac{n_4}{3}, \\
m_{33} &= 0.
\end{aligned}$$

Thus, it follows that the MLE solution is:

$$\begin{pmatrix} \hat{\theta}_N(mle) \\ \hat{\theta}_{1C}(mle) \\ \hat{\theta}_{2C}(mle) \end{pmatrix} = \mathbf{M}^{-1} \cdot \begin{pmatrix} n_2 \\ 0 \\ 0 \end{pmatrix}, \tag{1.14}$$

Along with $\hat{\theta}_{3C}(mle) = 1 - \hat{\theta}_{2C}(mle) - \hat{\theta}_{1C}(mle) - \hat{\theta}_{N}(mle)$.

See the Appendix for the 14 additional special cases and their solutions.

1.0.3 Monte Carlo Studies of the Model

Parameter Estimation with Monte Carlo

To ascertain the relative accuracy of the PPM and MLE methods for the EO model, a set of Monte Carlo simulations were conducted. Our Monte Carlo is sampling from a multinomial distribution to produce a set of values for frequencies that were generated from given values of the EO model parameters. Based on those frequencies, there were estimates of the EO model parameters' degree of accuracy.

This approach can be used to obtain parameter estimates from either categorical data or continuous data. Modifications of Monte Carlo samplings have also been employed for Bayesian parameter estimation (Seelig et al., 2020; DeCarlo, 2012), and for cases with either individual data or grouped data. In our study, we used Monte Carlo sampling to obtain parameter estimates for the EO model. We obtained parameter estimates for each condition and each trial of the memory task. This allowed us to determine whether there were any improvements in ordinal memory storage across tests for a given condition as well as differences across conditions.

MLE and PPM estimates of model parameters can differ from one another (Chechile, 2004, 2010b; Chechile et al., 2012). Simulation studies have provided evidence from Monte Carlo sampling experiments that PPM estimates are more accurate than corresponding MLE values. In the Monte Carlo studies, data were randomly sampled from known or stipulated population values for the underlying MPT parameters. Based on these frequencies, both MLE and PPM estimates were obtained. The mean absolute-value error over all the Monte Carlo simulations was found for both estimation methods. Although the mean absolute-value error decreased for both estimation methods as the sample size (*N*) increased, the PPM errors were uniformly smaller than MLE errors. Based on these results, we have to compare the relative accuracy of the PPM and MLE parameter estimation approaches for the EO model.

We collected 1,000 Monte Carlo samples to aid us in estimating the four storage parameters of the EO model, such that they all add up to 1. Collecting these samples was instrumental to assessing the error of both the MLE and PPM methods of parameter estimation. We knew the true population values for each sampled configuration, and based on such values, we produced a random sample of frequencies n_i, $i = \{1, \cdots, 4\}$ from a multinomial likelihood distribution. Based on these frequency values, we computed the MLE and PPM estimates of the EO model parameters. Then, for each parameter, the absolute value of the mean difference between the parameter estimate (PPM or MLE) and its true value was calculated. The resulting average errors are shown in Table 1 as a function of sample size. According to our results, we can see that as the sample size increases, the average parameter estimation error decreases. However, the MLE average errors are never smaller than those of the PPM method. Hence, the PPM estimation method yields more accurate results for our EO model. Previously, researchers have seen that the MLE method of parameter estimation is far from optimal, particularly for real-world experiments that use finite sample sizes. As an example, Le Cam observed that ”In view of Fisher’s vast influence, it is perhaps not surprising that the presumed superiority of the method is still for many an article of faith promoted with religious fervor. This state of affairs remains, in spite of a long accumulation of evidence to the effect that maximum likelihood estimates are often useless, or grossly misleading” (Cam, 1986).

Table 1 PPM (P) and MLE (M) Absolute-Value Errors as a function of n.

n	$\theta_{3C}(P)$	$\theta_{3C}(M)$	$\theta_{2C}(P)$	$\theta_{2C}(M)$	$\theta_{1C}(P)$	$\theta_{1C}(M)$	$\theta_N(P)$	$\theta_N(M)$
10	.140	.168	.161	.214	.153	.191	.125	.132
50	.073	.085	.087	.117	.086	.102	.059	.060
100	.052	.061	.074	.091	.065	.080	.044	.045
200	.037	.041	.049	.060	.048	.056	.029	.030
400	.029	.032	.039	.046	.037	.041	.022	.022
600	.023	.025	.032	.036	.028	.031	.017	.017
800	.020	.022	.028	.032	.025	.028	.015	.015
1000	.018	.019	.025	.028	.023	.025	.013	.013
2000	.013	.014	.019	.020	.016	.017	.009	.009

Note: Reprinted from Measuring components of the memory of order, by Chechile and Pintea (2021).

Pooling vs. Averaging

The nature of the data used to ultimately obtain MPT model parameter estimates is important. Whether we choose to use individual data and average the results across subjects or use pooled data for the group affects the relative accuracy of the model parameters. It has been shown that, in the context of MPT models, pooling leads to more accurate parameter estimates (Chechile, 2009, 2010a). The advantage of the pooling method has also been observed using the PPM and MLE approaches to parameter estimation (Chechile, 2009). Using PPM, pooling is consistently superior to averaging because of the use of the optimal posterior distribution: ”the pooling method is a natural way to obtain a final posterior distribution that is based on all the data, and it is also equivalent to an appropriately weighted mean of parameter estimates” (Chechile, 2009). For the MLE, depending on the parameter and the model, there may not be a pooling advantage. For these cases, the pooling and averaging yield the same results.

Chechile shows this in a proof of a theorem that delineates when pooling and averaging are equivalent (Chechile, 2009).

In our research, we pool the data we obtained from participants' performance on the AFC tasks in order to get the values of n_1, n_2, n_3, and n_4. This allows us to later calculate model parameters with the minimal average error. We anticipate that there will be differences between the two parameter estimation methods we use in this study (PPM vs. MLE). Furthermore, we predict that the PPM method will outperform the MLE.

Chechile (2009) addressed a basic question about the statistical analysis of MPT models concerning the relative accuracy of two different methods for estimating the parameters for a given experimental condition. One approach (the averaging method) is based on estimating the parameters for each participant separately and then averaging those values. The other approach (the pooling method) first groups the frequencies for each multinomial category across all the participants and then estimates the MPT model parameters. Importantly, the estimates for the EO model parameters do not satisfy this equivalence condition. Consequently the two methods of analysis will yield different values, and as a result, the question of the relative accuracy of the pooling and averaging methods is important to assess for the EO model.

To evaluate whether pooling is superior to averaging for the EO model, artificial Monte Carlo data can be generated for a given configuration of the four states of order knowledge. Given n_r observations per Monte Carlo pseudo-subject and given stipulated values for θ_{3C}, θ_{2C}, θ_{1C} and θ_N, an artificial set of data can be randomly sampled for n_g pseudo-subjects. These data can be examined by both the averaging and pooling methods. Because these data would have been generated from known values of the four order-knowledge probabilities, the accuracy of the two estimation methods can be compared. The overall accuracy of the estimates can be measured by the Kullback and Leibler divergence (Kullback and Leibler, 1951). That is, if θ_{3C}, θ_{2C}, θ_{1C} and θ_N are the stipulated population parameter values, whereas $\hat{\theta}_{3C}$, $\hat{\theta}_{2C}$, $\hat{\theta}_{1C}$ and $\hat{\theta}_N$ are the corresponding sample estimates, then the Kullback-Leibler divergence, d_{KL} is shown below. The superior estimation method is the one with the smaller value of value for d_{KL}.

$$d_{KL} = \theta_{3C} \ln \frac{\theta_{3C}}{\hat{\theta}_{3C}} + \theta_{2C} \ln \frac{\theta_{2C}}{\hat{\theta}_{2C}} + \theta_{1C} \ln \frac{\theta_{1C}}{\hat{\theta}_{1C}} + \theta_{N} \ln \frac{\theta_{N}}{\hat{\theta}_{N}} . \tag{1.15}$$

Our findings from a Monte Carlo study of the difference between averaging and pooling are displayed in Table 2. The true population parameters for these results were the following: 0.06 for θ_{3C}, 0.52 for θ_{2C}, 0.30 for θ_{1C} and 0.12 for θ_N. Across all cases, the Kullback-Leibler divergence measure is much smaller for the pooling method than for the averaging method. Our findings are consistent with those obtained for other MPT models explored by Chechile (2009).

While pooling is the more accurate method of parameter estimation for the EO model, there are cases where researchers have to obtain estimates on an individual-subject basis. One example is clinical studies, where individual measurement are often collected as part of the data. For settings such as this one, Chechile (2010a) recommended using the jackknifing method. According to this approach, jackknifed estimates are obtained for each individual. Each jackknifed value is a weighted difference between two pooled estimates of the parameter of interest. One estimate is based on pooling all data except for the individual *i*, while the other estimate is based on pooling all participants including participant *i*. Chechile (2010a) provides an overview and discussion of a study that used the jackknifed MPT estimation method for a group of Korsakoff amnesic patients vs. a yoked group of control patients. Overall, this approach can be used for the EO model if it is the case where parameter estimates have to be obtained on an individual basis.

Table 2 The Kullback-Leibler d_{KL} values for the pooling and averaging methods as a function of n_r observation per block and n_g Monte Carlo pseudo-subjects.

n_r	n_g	$d_{KL}(pooling)$	$d_{KL}(aver.)$
10	10	0.101	0.163
10	20	0.051	0.150
10	30	0.070	0.146
10	40	0.034	0.129
10	80	0.014	0.134
20	10	0.014	0.081
20	20	0.012	0.087
20	30	0.007	0.087
20	40	0.008	0.098
20	80	0.004	0.086
40	10	0.041	0.078
40	20	0.020	0.065
40	30	0.015	0.060
40	40	0.006	0.044
40	80	0.009	0.039

Note: Reprinted from Measuring components of the memory of order, by Chechile and Pintea (2021).

Chapter 2

Experiments 1 and 2

2.0.1 Experiment 1

Participants

The participants were 60 Tufts undergraduate students recruited using SONA Systems. Students signed a consent form prior to the beginning of the experiment received credit for a course requirement to participate in research or to write about a research project.

Design and Materials

Participants were subject to a memory task that consisted of a learning phase and a testing phase. The memory task was programmed using Python and was administered to participants on a lab computer. The program recorded the participant's data in real time, and it included information such as trial number, probe type, date of the session, start and end time of the session, correct attempts on the 6-AFC task, correct attempts on the 5-AFC task, correct attempts on the 4-AFC task, and incorrect attempts on the 4-AFC task.

Procedures

During the learning phase, participants were presented with a series of 40 words from one of six word-lists. The words appeared on the screen one at a time for two seconds each, following which the participants underwent trial 1 of the testing phase. After the presentation of the 40 words, the participants were given 6 separate tests of triplets – one triplet from each of the 6 within-subject conditions. Throughout trial 1, participants were presented with 6-AFC tasks that asked about the order of three words, which belonged to one of six possible conditions. We manipulated word location such that words were either Adjacent (right next to one another), Near (three words apart from one another), or Far (15 words apart from one another). We also manipulated word relatedness such that word triplets were either related by a category (e.g. apple, pear, orange) or completely unrelated to one another (e.g. star, trash, peninsula). We decided to manipulate relatedness because we believed this factor could have made a difference on participants' performance on recall of order information. In addition, we believed relatedness could have had theoretical significance for our understanding of order memory.

Given that word location has three levels and word relatedness has two levels, there are a total of six conditions that the participants are tested on. Following trial 1, participants reviewed the 40 words again before undergoing trial 2. After the end of trial 2, participants moved on to the next word-list until they got through all six. The entire session lasted approximately one hour, and a research assistant was present in the lab space at all times.

After obtaining the pooled frequencies for correct attempts on the 6-AFC task (n_1), 5-AFC task (n_2), and 4-AFC task (n_3), and for incorrect attempts on the 4-AFC task (n_4) (see Figure 5.1), we obtained estimates for the EO model parameters through Monte Carlo sampling using R. This sampling procedure created 50,000 Monte Carlo samples for the θ_{3C}, θ_{2C}, θ_{1C}, and θ_N parameters. The estimates are shown in Figures 5.2 and 5.3.

Results

The plots of the model parameter estimates and the statistical tests show the possibility of an interaction between word location and word relatedness. Figure 3 provides an overview of how order knowledge is distributed across the four EO model parameters across Trials 1 and 2. We can refer to Figure 7, which shows the pooled frequencies across conditions and trials. In addition, we can refer to Figure 8, which summarizes the average parameter estimates for the EO model for Related triplets, and to Figure 9, which does the same for Unrelated triplets. The data for Experiment 1 and all following experiments are documented in the Harvard Dataverse. See the Appendix for each experiment's website link.

Effect of Training. The effect of training trials was examined by grouping the frequencies across spacing and relatedness for both the first and second trials of the experiment. This way, we were able to use the PPM approach to estimate the EO model parameters for the two, separate levels of training. The result of this was $c_1 = 49,622$ coherent values for each of the EO model parameters that were sampled at random from the 50,000 mappings attempted for the first training trial. There was also an additional set of $c_2 = 49,999$ coherent mappings for each EO model parameter for the second training trial. Now, let $c^* = min(c_1, c_2)$. To verify whether the EO model parameters differed between the two training trials, we generated another set of vectors corresponding to difference scores for each of the model parameters. For context, the jth element for a difference score vector for a given EO model parameter corresponds to the difference between the jth element from the parameter vector for the second training trial and the jth element from the parameter vector for the first training trial. In the vector of difference scores, there are a total of c^* elements. According to our results, it appears that the Bayesian posterior probability that the parameter is larger for the second training trial than for the first is estimated by the proportion of difference scores greater than zero. Hence, the null hypothesis in this context is that the population difference score is $H_0 : \Delta \leq 0$, and the alternative hypothesis is $H_1 : \Delta > 0$. The Bayes factor BF_{10} is defined as follows:

$$BF_{10} = \frac{P(D|H_1)}{P(D|H_1)} = \frac{P(H_0)}{P(H_1)} \cdot \frac{P(H_1|D)}{P(H_0|D)}.$$

But because a flat prior distribution for Δ is assumed, it follows that the prior odds ratio for the two hypotheses is 1, which in turn means that $BF_{10} = \frac{P(H_1|D)}{P(H_0|D)}$. If $BF_{10} < 1$, then the Bayes factor $BF_{01} = \frac{P(D|H_0)}{P(D|H_1)} = \frac{1}{BF_{10}} > 1$ (Jeffreys, 1998; Kass and Raftery, 1995).

A Bayes factor greater than or equal to 19 (either BF_{10} or BF_{01}) is sufficiently strong evidence to warrant reporting as a reliable effect (i.e., it is a "good bet – too good to disregard"). A statistical hypothesis test in experimental science is best done in a context where both the null and the alternative hypotheses are credible (i.e., they are interval hypotheses rather than point hypotheses), and the researchers reports the Bayes factors. In this paper, all the hypotheses that are tested are interval hypotheses.

For testing the effect of learning trials, the θ_{3C} parameter reliably increases with training, $BF_{10} > 49,622$ (i.e., θ_{3C} is 0.226 on trial 1 and is 0.372 on trial 2). Neither the θ_{2C} nor the θ_{1C} parameters reliably changed with training. The respective Bayes factors were $BF_{10} = 2.06$ (θ_{2C} is 0.148 on trial 1 and is 0.162 on trial 2) and $BF_{10} = 2.74$ (θ_{1C} values being 0.059 and 0.078). Finally there is reliable decrease in the non-storage of order θ_N, $BF_{01} > 49,622$ for the change from 0.568 to 0.388.

Effect of Relatedness. The effect of word relatedness was analyzed using the same approaches we used to look at the effect of training. Furthermore, the results pertaining to relatedness were similar to those that we obtained when looking at training trials. When it comes to the θ_{3C} parameter, we found a reliable effect ($BF_{10} > 48,949$), since the average parameter estimates were 0.209 and 0.389 for the respective Unrelated and Related conditions. We did not find any reliable effects for either θ_{2C} ($BF_{01} = 2.34$, because we obtained average estimates of 0.163 and 0.146 for the Unrelated and Related stimuli) or for θ_{1C} ($BF_{10} = 6.94$, as we obtained average estimates of 0.050 and 0.087 for the Unrelated and Related conditions). However we did find a reliable effect for the θ_N parameter ($BF_{01} > 48,949$, since we got average estimates of 0.578 and 0.378 for the respective Unrelated and Related conditions).

Effects of Spacing. Because there are three spacing conditions, there are two different contrasts of interest. The first contrast is a comparison between the *Adjacent* and the *Far* conditions; whereas the second contrast is a comparison between the *Near* and *Far* separation conditions.

When comparing participant performance on the *Adjacent* vs. *Far* conditions, the Bayes factor associated with the θ_{3C} parameter had a value of $BF_{10} = 237.1$, which demonstrates that there is a high likelihood that this parameter is larger for the Adjacent condition (0.379) than for the Far condition (0.306). As for the θ_{2C} parameter, we obtained a Bayes factor of $BF_{01} = 107.2$, which shows that this parameter is more likely to be greater for the Far condition (0.203) than for the Adjacent condition (0.113). Finally, for both the θ_{1C} and θ_N parameters, we obtained small Bayes factors: $BF_{10} = 1.42$ and $BF_{10} = 1.57$, respectively.

For the comparison between the *Far* and *Near* conditions, the Bayes factor for θ_{3C} was $BF_{10} = 2,732.7$, which indicates that the parameter is larger (0.306) in the Far condition than it is in the Near condition (0.211). There was also a reliable effect for the θ_N parameter, $BF_{01} > 49,213$, indicating the non-storage is larger in the near condition (0.578) than in the Far condition (0.424). However, there was not a reliable difference for either the θ_{2C} parameter, $BF_{10} = 12.2$ or the θ_{1C} measure, $BF_{10} = 1.23$.

Overall Results for Experiment 1. The overall average of the $P(coh)$ statistic (associated with the PPM approach) for analyses we just reported has a value of 0.9927. As a result, we can infer that the posterior-to-prior odds ratio for the coherence of the EO model for this experiment is 11.92. Thus, the data from our experiment supports the EO model.

Figure 2.1 shows that Related words in the Adjacent condition (denoted by the label "AR") are clearly better remembered than those in the Near (NR) and Far (FR) conditions due to the greater amount of order information belonging to the θ_{3C} state. However, this is not true for Unrelated words. The Far (FU) words show a clear improvement between the first and second trials, unlike the Adjacent (AU) and Near condition (NU), which mostly remain the same across trials. We also see this pattern for the AR and FU words when looking at the percentage of order information in state θ_N. Related words in the AR condition have distinctly smaller amount of information

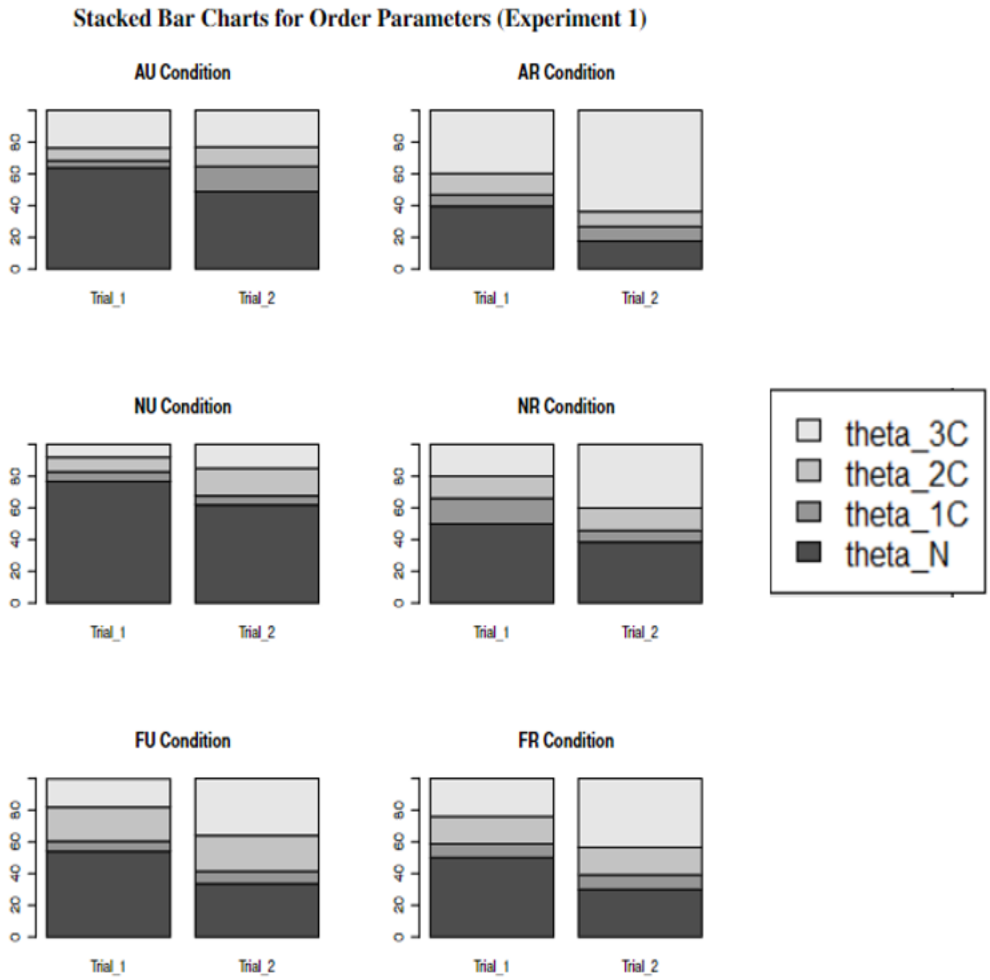

Figure 2.1: Stacked bar charts for the θ_{3C}, θ_{2C}, θ_{1C} and θ_N values (multiplied by 100) across test trials for each condition in Experiment 1. The different shades of gray represent the four states that order information can be in. The lightest shade of gray representing the θ_{3C} state. The next darker shade corresponds to the θ_{2C} state, followed by θ_{1C} and finally θ_N, the darkest shade of gray.

in the θ_N state than NR and FR words. Among Unrelated words, the words in the FU condition are the ones with the smallest θ_N values.

Discussion

The results show the possibility of an interaction between word location and word relatedness. Adjacent words are remembered best when they are Related, while Far words seem to be remembered better when they are Unrelated. In addition to this, Near words have the worst performance for both Related and Unrelated words.

We speculate that the order of Adjacent words is better remembered when they are related due to chunking. Chunking occurs when we break up a large amount of information into meaningful chunks, which help us effectively recall that information at a later time. Therefore, it is possible that the Adjacent-Related words are ”chunked” together, which could allow for better recall at the time of testing. We also hypothesize that the order of Far words is better remembered due to a temporal component of memory. It is possible that when words further away from each other, and thus closer to the beginning, middle, and end of the word-list, it is easier to identify the order the words appeared in. The distance in space and time of the Far words and the stark difference among them due to them being unrelated could be responsible for the better performance we see among Far-Unrelated words. The interaction is not altogether clear from this experiment alone. In addition to this, the learning of triplet order is still relatively low after two test trials. Therefore, participants need additional training to achieve higher levels of order knowledge, which we implement in Experiment 2.

2.0.2 Experiment 2

Participants

The participants were 117 Tufts undergraduate students recruited using SONA Systems. Students signed a consent form prior to the beginning of the experiment and received credit for a course requirement to participate in research or to write about a research project.

Design and Materials

The design and materials are similar to that of Experiment 1, except that the training of the list continued through four trials.

Procedures

During the learning phase, participants were presented with a series of 40 words from one of three word-lists. The words appeared on the screen one at a time for two seconds each, following which the participants underwent trial 1 of the testing phase. Throughout trial 1, participants were presented with 6-AFC tasks that asked about the order of three words, which belonged to one of the six conditions described in Experiment 1. Following trial 1, participants reviewed the 40 words again before undergoing trial 2. After the end of trial 2, participants reviewed the word-list again before undergoing trial 3. After trial 3, participants reviewed the word-list one last time before undergoing trial 4. After the end of trial 4, participants moved on to the next word-list until they got through all three. The entire session lasted approximately one hour, and a research assistant was present in the lab space at all times.

Results

We used the same procedure described in Experiment 1 to obtain the pooled frequencies for the four response categories (Figure 5.4). Figures 11 through 16 are tables of EO model parameters for the Adjacent-Related, Adjacent-Unrelated, Near-Related, Near-Unrelated, Far-Relted, and Far-Unrelated conditions, respectively. Our increase from two to four training trials helped us gather a clearer picture of the patterns in our data. All our data are available to the public through the Harvard Dataverse. See the Appendix for the website links corresponding to each experiment.

Effect of Training. For testing the effect of learning trials, we compared subject performance on trial 1 vs. trial 4. According to the data, the θ_{3C} parameter reliably increases with training, $BF_{10} > 40,865$ (i.e., θ_{3C} is 0.1445 on trial 1 and is 0.4819 on trial 4). There was no reliable increase for θ_{2C}. Its average was 0.1949 for trial 1 and 0.2042 for trial 4, where $BF_{10} = 1.637$. In addition, there was no reliable increase for θ_{1C}, which had an average value of 0.0308 at trial 1 and an average value of 0.0338 by trial 4 ($BF_{10} = 1.2434$). Finally there is a reliable decrease in the non-storage of order θ_N, $BF_{01} > 40,865$ for the change from 0.6299 to 0.2801.

Effect of Relatedness. Our findings demonstrate the same pattern we previously saw in our first experiment: participants performed remarkably better on items in the Related condition. When looking at the θ_{3C} parameter, we found a reliable effect

($BF_{10} > 47,760$), as we found that the average values for θ_{3C} were 0.4174 and 0.2659 for the Related and Unrelated conditions, respectively. Furthermore, we found a reliable effect for θ_{2C} ($BF_{01} = 53.96$), as the average parameter values we obtained were 0.1750 and 0.2208 for the Related and Unrelated items, respectively. We did not, however, find a reliable effect for θ_{1C}: we obtained average parameter values of 0.0565 and 0.0283 for the Related and Unrelated items ($BF_{10} = 10.69$). Lastly, we found a reliable change for θ_N: we obtained average parameter values of 0.3510 and 0.4850 for the Related and Unrelated conditions, respectively ($BF_{01} > 47,760$).

Effects of Spacing. For the comparison between *Adjacent* and Far, the Bayes factor for the θ_{3C} measure was $BF_{10} > 20,750$, indicating that it is highly likely that the parameter is larger in the Adjacent condition (0.4275) than the Far condition (0.3207), which we saw in experiment 1. The corresponding Bayes factor for the θ_{2C} parameter was $BF_{01} > 20,750$, indicating that the parameter is smaller in the Adjacent condition (0.1597) than in the Far condition (0.2916), which we also saw in Experiment 1. There was also an effect for the θ_{1C}. Its corresponding Bayes factor was $BF_{01} = 31.02$, showing that it is highly likely that the parameter is larger for the Far condition than for the Adjacent condition. Finally, we also detected an effect for θ_N, with $BF_{10} = 1,210.59$. This suggests that it is highly likely that this parameter is larger for the Adjacent condition than for the Far condition.

When we compared the *Far* and *Near* conditions, we only detected a reliable effect for θ_{2C} ($BF_{01} > 49,933$), which shows that the average value of this parameter is more likely to be larger in the Far condition (0.2916) than in the Near condition (0.1473). As for θ_{3C}, the corresponding Bayes factor ($BF_{10} = 0.0703$) showed that there was no difference between the Near and Far conditions. Furthermore, we did not find a reliable effect for θ_{1C} ($BF_{10} = 3.074$). However, we did find an effect for θ_N ($BF_{10} > 49,933$), which allowed us to conclude that the non-storage parameter is more likely to be larger in the Near condition (0.4949) than in the Far condition (0.3378).

Figure 2.2, which is a stacked-bar chart display of all the conditions, shows that Related words in the AR condition are still better remembered than those in the NR and FR conditions. As for Unrelated words, the order memory for words in the FU condition is the best for both Trials 1 and 2. The words in the AU condition perform as poorly as the NU condition in trial 1, but by trial 2 the two conditions pull apart and the AU words perform better than the NU ones, since the amount of information in the θ_N state is still greater for words in the NU condition.

Finally, if we pay close attention to amount of order information in the θ_N state, we get a picture that is highly similar to the one we obtained in Experiment 1. When it comes to Related words, the amount of information in the θ_N state is the smallest for the AR condition. Notably, words in the FR and NR conditions have the same performance for trial 1, but by trial 2 the words in the Far condition perform better and pull away from the words in the Near condition (the percentage of information in the θ_N state in the NR condition is still greater than that of the FR condition). As for Unrelated words, words in the FU condition have the smallest amount of order information in the θ_N state, followed by the AU, and NU words.

Discussion

According to the results, we have a clear interaction between word location and word relatedness. As we inferred in Experiment 1, Adjacent words perform best when they are Related, while Far words perform better when they are Unrelated. Words in the Near condition consistently have the worst performance regardless of word relatedness. The clearer and consistent results we obtained in this experiment support our previous speculations in Experiment 1 that chunking for AR words and temporal memory for FU words are responsible for the patters we observed.

Despite this, the plots show that participants struggle to learn word order in the memory task. Looking at Figure 2.2, by the end of the experiment, the amount of order information for NR and FR words that reaches the θ_{3C} state is about 50%. Among Unrelated words, participants perform even worse: the amount of order information that reaches the θ_{3C} state for NU and FU words is slightly below 40%. Subjects struggle with learning the order of all three items when they are Near and Far.

Stacked Bar Charts for Order Parameters (Experiment 2)

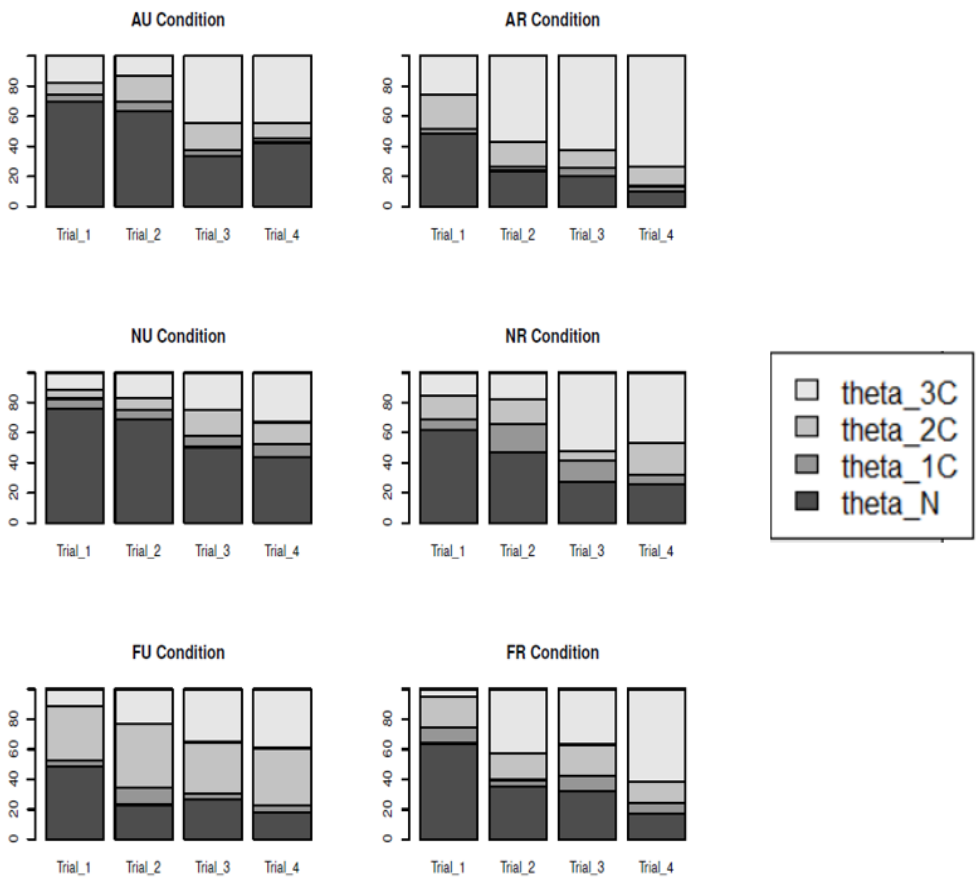

Figure 2.2: Stacked bar charts for the θ_{3C}, θ_{2C}, θ_{1C} and θ_N values (multiplied by 100) across test trials for each condition in Experiment 2.

Chapter 3

Experiment 3 and Item vs. Order Learning

We decided to incorporate item generation as part of the learning phase in Experiment 3, in hopes that it will improve order memory in participants. The generation effect in the memory literature refers to the enhanced learning that occurs when one has to produce or complete the target items (see Chechile 2018 for an extensive review of generative encoding procedures and effects). Therefore, we designed the procedure for our third experiment to have more of a generative task.

We have noticed that Related words are generally better remembered than Unrelated words. Therefore, we only focus on Unrelated words in Experiment 3 and leave out the Adjacent condition in order to gather more information solely focusing on words in the Near and Far conditions. Lastly, we chose to incorporate item-content questions as part of the testing phase to gather data on their memory for word content in the same AFC format as the order-knowledge questions. These procedures are explained in further detail in the next section.

3.0.1 Item-Storage (IS) Model

In addition to looking at the four stages of order memory storage, Experiment 3 aims to investigate the stages of item content storage. We use pooled frequencies from

Multinomial Processing Tree Model for Item Storage

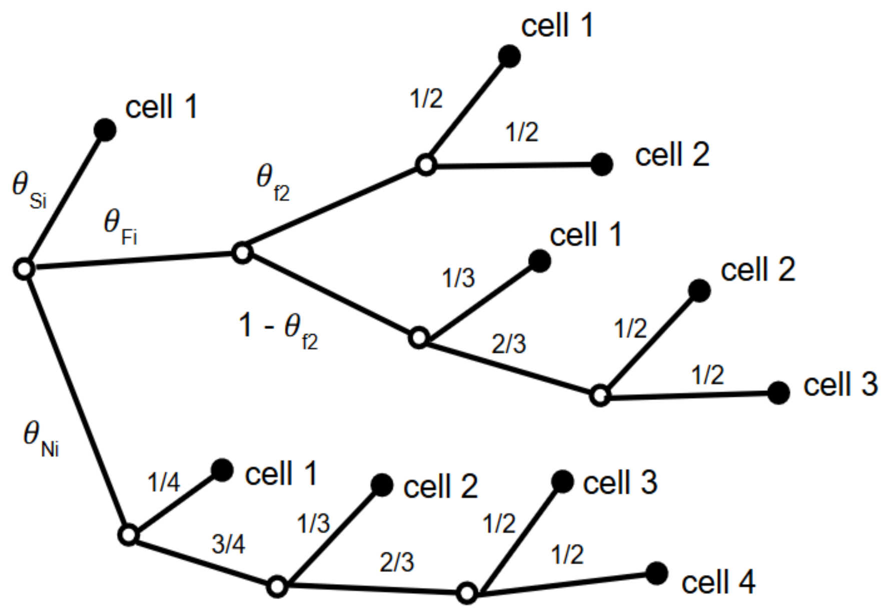

Figure 3.1: Multinomial processing tree model for item storage. The illustration above follows the same structure as the MPT model we constructed for the EO model.

item-content questions to obtain parameter estimates for an item-storage (IS) model (Figure 5.14). We define the following levels for item storage: correct attempts on the 4-AFC task (sufficient item storage - θ_{Si}), correct attempts on the 3-AFC task (fractional item storage - θ_{Fi}), correct attempts on the 2-AFC task (fractional item storage - θ_{Fi2}), and incorrect attempts on the 2-AFC task (no item storage - θ_{Ni}, where $\theta_{Si} + \theta_{Fi} + \theta_{Ni} = 1$. We obtained the estimates for the IS model through Monte Carlo sampling with R, which created 50,000 Monte Carlo samples for θ_{Si}, θ_{Fi}, θ_{Fi2}, and θ_{Ni}. A visual representation of the IS model is shown in Figure 3.1.

There are two cases for partial storage. There is partial storage if the participant can reject either one (θ_{Fi}) or two (θ_{f2}) of the three foil items on the 4-AFC test. If the participants cannot reject any of the foils, then it is corresponds to the case represented by non-storage (θ_{Ni}). If the person can reject all three foils on the 4-AFC test, then it corresponds to the case represented by sufficient storage (θ_{Si}). Because there are two types of fractional storage, there is a parameter θ_{f2} to represent the conditional probability for the rejection of two foils when there is partial item knowledge. Thus, the conditional probability for the rejection of only one of the foils is $1 - \theta_{f2}$ for the three foils. The population cell proportions are the following:

$$
\begin{aligned}
\phi_{1(i)} &= \theta_{Si} + \frac{1}{2}\theta_{Fi}\theta_{f2} + \frac{1}{3}\theta_{Fi}(1-\theta_{f2}) + \frac{1}{4}\theta_{Ni}\,, && (3.1)\\
\phi_{2(i)} &= \frac{1}{2}\theta_{Fi}\theta_{f2} + \frac{1}{3}\theta_{Fi}(1-\theta_{f2}) + \frac{1}{4}\theta_{Ni}\,, && (3.2)\\
\phi_{3(i)} &= \frac{1}{3}\theta_{Fi}(1-\theta_{f2}) + \frac{1}{4}\theta_{Ni}\,. && (3.3)\\
\phi_{4(i)} &= \frac{1}{4}\theta_{Ni}\,. && (3.4)
\end{aligned}
$$

With the PPM estimation procedure random samples from ϕ-space mapped to a corresponding vector of the IS model parameter values, provided that the mapping is coherent. For the uth sample from ϕ-space, denoted as the vector $\langle \phi_{1(i)(u)}, \cdots, \phi_{4(i)(u)} \rangle$, the corresponding values for the item-storage parameter are:

$$
\begin{aligned}
\theta_{Si}(u) &= \phi_{1(i)}(u) - \phi_{2(i)}(u)\,, && (3.5)\\
\theta_{Fi}(u) &= 2\phi_{2(i)}(u) + \phi_{3(i)}(u) - 3\phi_{4(i)}(u)\,, && (3.6)\\
\theta_{Ni}(u) &= 4\phi_{4(i)}(u)\,, && (3.7)\\
\theta_{f2}(u) &= \frac{2[\phi_{2(i)}(u) - \phi_{3(i)}(u)]}{2\phi_{2(i)}(u) + \phi_{3(i)}(u) - 3\phi_{4(i)}(u)}\,, && (3.8)
\end{aligned}
$$

Provided that four constraints are satisfied for a coherent mapping. These constraints are:

$$
\begin{aligned}
\phi_{1(i)}(u) &\geq \phi_{2(i)}(u)\,, && (3.9)\\
\phi_{2(i)}(u) &\geq \phi_{3(i)}(u)\,, && (3.10)\\
\phi_{3(i)}(u) &\geq \phi_{4(i)}(u)\,, && (3.11)\\
\phi_{4(i)}(u) &\leq \frac{1}{4}\,. && (3.12)
\end{aligned}
$$

If the condition stipulated in equation (3.9) is violated, then that would imply the incoherent result of $\theta_{Si} < 0$. If the condition specified in equation (3.10) is violated, then would imply the incoherent result of $\theta_{f2} < 0$. If the requirement set out in equation (3.11) is violated, then that would imply that $\theta_{f2} > 1$. Finally if the equation (3.12) constraint is violated, then it would imply that $\theta_{Ni} > 1$. As with the EO model, only the coherent Monte Carlo samples from the posterior Dirichlet distribution are used for the estimation of the parameters for the IS model.

3.0.2 Experiment 3

Participants

The participants were 50 Tufts undergraduate students recruited using SONA Systems. Students signed a consent form prior to the beginning of the experiment and received credit for a course requirement to participate in research or to write about a research project. Though the number of participants here is lower than the previous two experiments, we were able to obtain parameter estimates with standard deviations below 0.05. Had it not been for the sudden halt on research that the pandemic brought forth in early 2020, we would have run more participants.

Design and Materials

Participants were subject to a memory task that was designed and implemented in the same way as in Experiments 1 and 2.

Procedures

During the learning phase, participants were presented with a series of 32 words from one of four word-lists. The words appeared on the screen one at a time, each with one missing vowel. Participants were instructed to press the letter on the keyboard that corresponded to the missing letter, after which both the complete and incomplete word were displayed on the screen for 2.5 seconds before moving on to the next word. We chose this manipulation for the generative task because previous studies that looked at generation effects used this same manipulation as well (e.g. Chechile & Soraci, 1999), where items had missing letters during the learning phase.

Participants then underwent trial 1 of the testing phase. Throughout trial 1, participants were presented with 6-AFC tasks that asked about the order of three words, which belonged to one of two possible conditions. We manipulated word location such that words were either Near (two words apart from one another) or Far (11 words apart from one another). All words were unrelated to one another. In addition to these conditions, participants were presented with item-content 4-AFC tasks that prompted them to select the word they recognized seeing in the learning phase. The remaining three options were foil items that were never presented in the learning phase. Gathering data on item content and order knowledge allowed us to decouple the two and obtain a clearer picture of how participant performance differs between the two.

Following trial 1, participants reviewed the 32 words again before undergoing trial 2. After the end of trial 2, participants reviewed the word-list again before undergoing trial 3. After trial 3, participants reviewed the word-list one last time before undergoing trial 4. After trial 4, participants moved on to the next word-list until they got through all four. The entire session lasted approximately one hour, and a research assistant was present in the lab space at all times.

Results

All of our data are available through the Harvard Dataverse. See the Appendix for the website links corresponding to each experiment. Figures 18 and 19 are tables of the frequencies for the EO task and model for the respective Near and Far conditions. The frequencies for the item storage (IS) model are shown in Figure 17, and the IS parameter estimates are shown in Figure 20.

Effects of Spacing. For the θ_{3C} parameter, after the first training trial, there is a reliably higher value in the *Far* condition (0.196) than in the *Near* condition (0.047); the Bayes factor for this test is $BF_{10} = 89.3$. There is also a reliable difference between the far and near conditions for the θ_{2C} ($BF_{10} = 34.3$), but there is not a reliable difference between these conditions for the θ_{1C} parameter ($BF_{01} = 2.82$). For the θ_N parameter, there is a reliably higher value in the near condition rather than in the far condition ($BF_{01} = 4,080$). However, some of these spacing effects are different after four training trials. For the θ_{3C} parameter there is not a reliable spacing effect ($BF_{10} = 7.85$) after four trials. Also, the θ_{2C} and θ_{1C} parameters did not reliably vary with spacing; the respective Bayes factor values are $BF_{10} = 10.3$ and $BF_{01} = 9.4$ for those tests. However, the θ_N parameter remained reliably higher in the near condition ($BF_{01} = 4,080$).

Item Storage. When it comes to item storage, the expression $(1-\theta_{Ni})^3$ represents the probability of some item knowledge of a triplet. So, the null hypothesis that order and item knowledge is not different would be $(1-\theta_N) \geq (1-\theta_{Ni})^3$. Alternatively, we can have H_0 expressed as $\theta_N \leq 1-(1-\theta_{Ni})^3$. Overall, we found item storage to be at a distinctly higher level than order knowledge. After trial 1, the estimate non-storage for an item was $\hat{\theta}_N = 0.0715$. Hence $(1-\hat{\theta}_{Ni})^3 = 0.8003$. Consequently the null hypothesis is $H_0 : \theta_N \leq 0.1997$, and the alternative hypothesis is $H_1 : \theta_N > 0.1997$. Using the values in the PPM θ_N vector results in strong evidence for the alternative hypothesis (i.e., $BF_{10} > 29,380$). Consequently after a single training trial, the knowledge gained about items is reliability higher than the knowledge of the ordinal position of the items. This same hypothesis can also be examined after the fourth training trial. Again the Bayes factor strongly favored the alternative hypothesis (i.e., $BF_{10} > 35,730$). To learn the order of items in memory requires more than just storing the items. Information about ordinal position also needs to be encoded. Consequently, these statistical conclusions, done by using MPT measurement, supports the long standing claim by Underwood (1977) that ordinal information is a different attribute of memory.

The parameter estimates for the EO model show that words in the Far condition are still performing better than those in the Near condition, as we previously observed in Experiments 1 and 2. However, the inclusion of the generation effect did not boost performance on order memory at all. The percentage of order information in the θ_{3C} state for Far words reached about 33%, and for Near words it reached 24%.

Stacked Bar Charts for Order Parameters (Experiment 3)

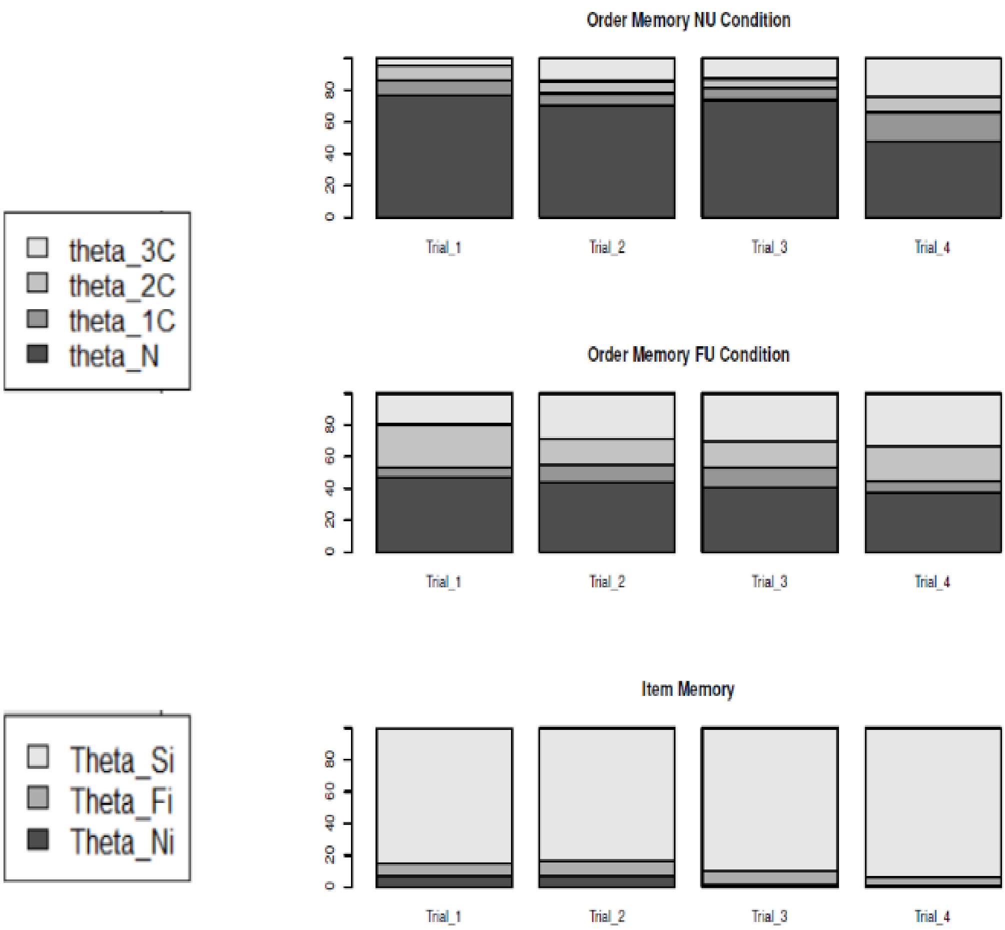

Figure 3.2: Stacked bar charts for the θ_{3C}, θ_{2C}, θ_{1C} and θ_N values (multiplied by 100) across test trials for each condition in Experiment 3.

As for the parameter estimates of the IS model, we see that item storage is consistently high. On average, for a given word-list, the θ_{Si} value increases from 85% to 94%. The corresponding values for fractional storage (θ_{Fi} and θ_{Fi2}) as well as that of no item storage (θ_{Ni}) are close to zero.

Discussion

The results from both the EO model and the IS model show evidence that item-sufficient storage is very high while knowledge of order memory is not as high. We can infer that item knowledge and order knowledge are separate attributes of memory that have a different rate of learning. Furthermore, despite our implementation of a generative task, participants still had lots of room for improvement for order knowledge. We speculate that participants may either require additional repetitions, or they may benefit from a different generative task.

3.0.3 Discussion of First Three Experiments

A major goal for this paper was to establish our method for investigating order memory and show that it extracts measures for four levels of order knowledge about a triplet of three probed items for each condition. The four measures are estimates of probabilities. How are these values supposed to be interpreted? For a given participant on a single trial, there are four states of knowledge for a triplet of items. The participant's knowledge is either (1) complete about the order relationship of all three items, (2) missing one order relationship, (3) missing two order relationships, or (4) no knowledge of order at all. By grouping over participants or grouping over similar test trials for the same participant, the MPT model results in probability values for these four states. The most straightforward way to interpret the probability values is as proportions for the above four knowledge states. In essence, the MPT model is a tool to detect and quantify mixture rates. The high construct validity for our model is due to the series of AFC tasks; without the follow-up testing when there are errors, there would not be sufficient data to enable a decomposition of the mixture over the four possible states of knowledge. The forced-choice method also has the advantage of establishing the guessing rates without requiring free model parameters to deal with guessing.

The results from the three experiments show evidence that item memory and order memory are different. In Experiment 1, we saw that there was a possibility for an interaction between word location and word relatedness. In Experiment 2, we verified this with a larger sample and additional trials and obtained results consistent with the interaction we observed in Experiment 1. And finally, in Experiment 3, we were able to identify a difference between the rate of learning of item content and item order. We can thus infer from these results that item knowledge and order knowledge are distinct attributes of memory.

There are several hypotheses and models in the psychology literature that were generated for the purpose of describing participants' ability to order a series of stimuli. In light of our findings, these hypotheses and models are not the most appropriate. As an example, let us consider strength theory (e.g. Hinrichs (1970)). According to this hypothesis, judging an item's recency is a function of that item's relative strength: stronger events are assumed to be the more recent ones. It is assumed that strength is developed based on the initial encoding during the learning phase, and it decays as time passes. It all learned items have the same amount of initial strength during the encoding process, then it is assumed that the stronger memory trace will correspond to the more recent trace, since the amount decayed would be smaller for recent items. Furthermore, according to strength theory, the difficulty in recency judgment is minimized with increased item spacing, because greater spacing allows for more distinguishable strength values. As a result, one would expect the items in the adjacent condition to have the worst performance, while items in the far condition are expected to have the best performance. We know, however, that order knowledge is superior for items in the adjacent condition, as it was associated with better order knowledge than the near and far conditions. Thus, the prediction provided by the strength theory does not match the results we obtained from our first two experiments. In addition, strength theory does not account for our findings from our third experiment: by relying on strength value alone during the encoding process, how can this theory account for the difference between the item content knowledge and the order knowledge that is encoded? Hence, due to the challenges it faces, the strength hypothesis does not make an appropriate account for order memory.

The reconstructive hypothesis (Friedman, 1993) is another problematic theoretical position that rejects the idea that time or order information is stored in memory. According to Friedman, human reasoning can be used to guide recollection about time and order in autobiographical memory (1993, p. 47). Though we can agree with this, this process by itself cannot account for the four states of order knowledge measures with the EO model. Moreover, a reconstructive process is also not a creditable mechanism for accounting for how the probability values of the four knowledge states change as a function of the independent variables examined in the experiments.

Another hypothesis that does not hold for order learning is the associative chaining model (Ebbinghaus, 2013). This particular account does not include or require the temporal and ordinal aspects of an event to be stored, which is reminiscent of the strength model and of the reconstructive model. The events are linked in an associative chain $A - B - C -$ .... However, there is an issue with relying on bidirectional associative bonding to encode order knowledge. Therefore, a better way of characterizing the three events would be A $\rightarrow$ B $\rightarrow$ C, where the symbol $\rightarrow$ represents the flow of time. Changing the way we look at how the three items are learned leads to an important theoretical shift: we now see that the link between the three stimuli includes ordinal and temporal information, and it is no longer a simple association between two events.

Chaining theory aims to rely on bidirectional strength linkage between events to account for time. The problem with this, however, is that a bidirectional link only accounts for the association between the two events. We need something more outside of this link in order to account for ordinal information. However, even with the change we made with our notation, the chain architecture is still flawed. First, the architecture of a linked chain does not account for partial order knowledge, in which the individual knows two of the three order relationships. With the way the chaining model arrangement is set up, we can only account for either of the following: complete order knowledge (both links are intact at the time of test); one order relationship (if only one link is intact); or no order knowledge whatsoever (neither of the two links is intact). However, our experiments thus far show that one can be in the partial knowledge state, where one has knowledge of two relationships among the three events. Without the aid of an MPT model, the associative-chaining theory can only claim that the data about the knowledge of two of three is due to guessing (i.e. one relationship is known while there is a lucky guess for the order relationship). However, our EO model already takes guessing into account. We have shown in the first three experiments that the probability interval for the θ_{2C} state does not include zero. This result is inconsistent with the chaining theory. As a result, chaining theory requires modifications to account for this problem.

The fact that the hypotheses aforementioned fail to account for our results has important theoretical consequences. It appears that either time or order information has to be encoded in memory: memory storage has to include content information related to the learned event along with temporal information. Our results from the EO model corroborate Underwood's claims (1969, 1977) that ordinal information is a distinct attribute of memory. However, Underwood did not create models to describe memory processes. As a result, he did not introduce any specific theories about how ordinal information was encoded, learned, accessed, and altered. There are several theoretical hypotheses about memory encoding either time or order along with item content information (Brown et al., 2007, 2000; Chechile, 2018; Howard et al., 2015; Howard and Eichenbaum, 2015; Ward et al., 2013). Using our EO task and model to investigate order memory has implications for these theories, as they would need to be able to explain our findings. For instance, they would have to explain the following: (1) order learning being slower than item content learning; (2) order learning not being an all-or-none phenomenon for a triplet of events and having four states of order knowledge; (3) repetition of training improving order knowledge; (4) spacing of events affecting the quality of the encoded order information; finally, (5) item relatedness improving order learning. It is still not yet clear whether any of the models and theories aforementioned can account for all these findings in our present research.

Chapter 4

Experiment 4 and Overall Discussion

Thus far, the three experiments we conducted show us differences in learning triplets of words that are related and unrelated. We also know that there is a stark difference in the learning rate of order versus item content. However, there is one pattern we have consistently seen throughout our experiments: despite the number of trials we've used, order learning is still not as good as it could be. As a result, one of our goals in our forthcoming work is to double the number of trials (from four to eight) in hopes that participants can reach better results with their order learning.

In the research study we propose here, we plan to continue looking at relatedness. In addition, it would be interesting to characterize this learning process through a different approach. We plan to use a Markovian process to explain the changes in order knowledge as participants progress through the AFC tasks. Chechile and Sloboda (2014) have previously advocated for the use of Markovian processes in to describe memory processes, which further motivates our use of this approach.

4.1 Experiment 4

4.1.1 Participants

The participants were 81 undergraduate Tufts students recruited online through SONA Systems. Each participant was screened to be at least 18 years of age, and they consented to participating in the study prior to the beginning of the experiment. All participants received course credit for participation following the conclusion of the experiment.

4.1.2 Design and Materials

Participants were subject to a memory task that follows the same design as the task we conducted in our first three experiments. The memory task was implemented using Python along with the PsychoPy package. The program recorded the participant's data in real time, and it included information such as trial number, probe type, date of the session, start and end time of the session, correct attempts on the 6-AFC task, correct attempts on the 5-AFC task, correct attempts on the 4-AFC task, and incorrect attempts on the 4-AFC task. The program also recorded the information of the responses on tests of item knowledge.

4.1.3 Procedure

During the learning phase, participants were presented with a series of 40 words from one of two word-lists. The words appeared on the screen one at a time for 2.5 seconds each, following which was followed by the trial 1 test phase. Half of the word triplets were related items (e.g., apple, pear, orange) and the other half of the triplets were unrelated items (e.g., star, trash, peninsula). Test trials also included item-content questions. Thus, after each training trial there were two tests for related triplets, two tests for unrelated triplets, and two tests for item knowledge. Participants were tested two times in: order memory for related triplets, order memory for unrelated triplets, and item memory. All stimuli, lists, and triplets tested for each condition are provided in the Appendix. Each of the eight training trials had the same presentation and testing protocol. A second list was used for another eight training and testing trials.

4.1.4 Results

Section I (Grouping All Participants)

We first organized the participants' responses for order learning into pooled frequencies (n_1, n_2, n_3, n_4) in the same fashion that we did in our previous three experiments. Table 3 below displays these pooled frequencies and how they differ by condition. It is evident that learning is occurring in both conditions, as n_1 increases and n_4 decreases for both Unrelated and Related triplets. We followed the same process for item learning, as shown in Table 4. We can already see how starkly larger the values for n_1 are for item learning compared to those of order learning in Table 3.

Table 3 Pooled Frequencies for Order Learning of Unrelated and Related Items.

	Unrelated				Related			
Trial	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4
1	101	66	44	113	152	61	43	68
2	118	69	50	87	152	69	39	64
3	155	63	30	76	214	50	20	40
4	152	79	34	59	211	56	29	28
5	181	61	26	56	229	51	11	33
6	200	58	23	43	245	39	10	30
7	186	67	24	47	246	43	15	20
8	227	49	20	28	254	45	8	18

Table 4 Pooled Frequencies for Item Learning.

Trial	n_1	n_2	n_3	n_4
1	246	36	25	17
2	246	48	22	8
3	290	16	13	5
4	298	11	11	4
5	285	24	12	3
6	293	19	7	5
7	310	9	5	0
8	309	10	1	4

However, we have to take a better look at how the learning rates for order differ between conditions. We calculated the EO model parameters for Unrelated and Related triads, which are displayed in Table 5. According to this table, participants had an easier time learning Related items (θ_{3C} = 0.640) over Unrelated items (θ_{3C} = 0.543) by the time they finished the memory task. As for item learning, we can turn to Table 6, which shows that item learning was drastically easier for participants to acquire over the course of the experiment (θ_{Si} = 0.910).

Table 5 EO Parameters for All Subjects (Unrelated vs. Related).

	Unrelated				Related			
Trial	θ_{3C}	θ_{2C}	θ_{1C}	θ_N	θ_{3C}	θ_{2C}	θ_{1C}	θ_N
1	0.107(0.038)	0.126(0.057)	0.083(0.053)	0.684(0.048)	0.270(0.042)	0.122(0.058)	0.182(0.062)	0.426(0.045)
2	0.148(0.041)	0.122(0.061)	0.194(0.066)	0.536(0.049)	0.256(0.043)	0.173(0.062)	0.170(0.062)	0.401(0.044)
3	0.281(0.042)	0.192(0.056)	0.068(0.046)	0.460(0.043)	0.500(0.042)	0.188(0.050)	0.068(0.040)	0.243(0.035)
4	0.223(0.045)	0.274(0.063)	0.138(0.057)	0.365(0.042)	0.475(0.043)	0.161(0.056)	0.191(0.052)	0.173(0.031)
5	0.366(0.043)	0.210(0.055)	0.081(0.047)	0.344(0.040)	0.550(0.041)	0.221(0.046)	0.032(0.026)	0.197(0.030)
6	0.433(0.043)	0.212(0.054)	0.089(0.046)	0.267(0.037)	0.639(0.038)	0.158(0.040)	0.027(0.023)	0.176(0.028)
7	0.363(0.045)	0.260(0.056)	0.087(0.047)	0.291(0.038)	0.623(0.039)	0.172(0.046)	0.075(0.037)	0.129(0.027)
8	0.543(0.041)	0.177(0.045)	0.104(0.044)	0.177(0.031)	0.640(0.039)	0.221(0.042)	0.030(0.022)	0.110(0.023)

Table 6 IS Model Parameters (All Subjects).

Trial	θ_{Si}	θ_{Fi}	θ_{Ni}
1	0.634(0.037)	0.154(0.050)	0.212(0.044)
2	0.608(0.040)	0.282(0.050)	0.110(0.035)
3	0.829(0.028)	0.100(0.033)	0.071(0.027)
4	0.865(0.024)	0.075(0.029)	0.058(0.023)
5	0.793(0.031)	0.159(0.036)	0.048(0.023)
6	0.839(0.028)	0.098(0.031)	0.063(0.023)
7	0.915(0.021)	0.073(0.022)	0.012(0.011)
8	0.910(0.021)	0.054(0.021)	0.036(0.014)

Our initial attempt to model the learning of order and the learning of items was to examine the Bower (1961) model, which has two states (i.e., learned and unlearned). This model assumes on each trial that there is a 2 by 2 transition matrix with the same values, which are $a_{11} = 1$, $a_{12} = 0$, a_{21}=β and $a_{22} = 1 - \beta$. The upshot of this assumption, the model predicts that a learned item on trial n will remain learned on trial $n+1$, and an item unlearned on trial n has a probability equal to beta for being learned on trial $n+1$. A key assumption is beta remains fixed as a function of trials. This is the stationarity assumption of a Markov chain.

Although we have already shown that learning for both order and item have partial learning, the two-state model is nonetheless useful. In our case, the learned state is grouping the complete knowledge and partial knowledge states into a common state. Chechile and Sloboda (2014) also showed that the beta parameter in the Bower model was a learning hazard rate. Learning hazard is increment in the probability of learning at trial n given that it was not learning previously. So, the Bower model implicitly assumes that the learning hazard is a constant.

In this Markov chain, the probability for information to be in the non-storage state after n trial is $(1-\beta)^n$. Note that a key assumption of any Markov chain is that the coefficients in the transition matrix remain constant. Importantly, we have a measurement of non-storage for both order triplets and item knowledge. For example, for the EO model the θ_N parameter on the n-th trial should be equal to $(1-\beta)^n$. Consequently, the θ_N parameter per trial provides a means to estimate the β parameter. That parameter should be constant. See Table 7 for the β values and their 95% probability intervals for both the Related and the Unrelated conditions. Clearly, the stationarity assumption of a constant beta is not valid.[1]

Table 7 β_N Values for Related vs. Unrelated Triplets

Trial	$\beta_{N(Related)}$	$\beta_{N(Unrelated)}$
1	0.574 [0.488, 0.663]	0.316 [0.224, 0.410]
2	0.367 [0.304, 0.441]	0.268 [0.203, 0.334]
3	0.376 [0.317, 0.433]	0.228 [0.184, 0.277]
4	0.355 [0.299, 0.411]	0.223 [0.180, 0.269]
5	0.277 [0.240, 0.325]	0.192 [0.158, 0.231]
6	0.251 [0.214, 0.293]	0.198 [0.164, 0.237]
7	0.254 [0.214, 0.303]	0.162 [0.134, 0.195]
8	0.241 [0.205, 0.284]	0.195 [0.163, 0.233]

Chechile and Sloboda (2014) showed that the β value in the Bower model was a learning hazard. Since it was assumed to be a Markov Chain, it was assumed that the β value would be the same for all trials. However, we established that the β values are systematically decreasing (see Table 7). For the Related condition, the β_1 value (0.574) was reliably greater than that of trials 2 through 8. In addition, the value of β_4 was also reliably different from that of trials 5 through 8. For the Unrelated condition, the β_1 and β_2 values were statistically larger than that of trials 5 through 8. To comple-

[1]The linear operator model from Bush and Mosteller (1951) also has a fixed learning parameter, so it also cannot be an accurate model of order learning.

ment this information, we obtained correlation coefficient values for the relationship between the trial number and the β values for each condition. We obtained a Kendall τ value of -0.8571 for the Related β values, which tells us that there is a strong, negative correlation between the two quantities. A Bayesian assessment of the Kendall tau rank-based correlation (Chechile, 2020) shows that the beta values decreased reliability for the Related condition (.9999992) as well as for the Unrelated condition (.9999924).

Given our results, we need a theory that can account for non-constant hazard. Chechile and Sloboda also showed that the Weibull model for learning is flexible enough to represent the learning procedure as being either monotonically decreasing hazard (slower learning with trials), constant hazard (like the Bower model), or monotonically increasing hazard (faster learning with trials).

We begin obtaining the Weibull fits starting with the θ_{3C} parameter, as shown in Table 8. With our EO parameters from the memory task, we can conduct a Weibull model fitting for each parameter with the following steps. We begin this fitting method with the following equation.

$$\theta_{3C} = 1 - e^{-a_3 t_n^c}$$

Where a_3 is a fitting parameter for the group we are looking at. Both a_3 and c are constants, and t_n represents the trial number. The a_3 and c parameters are both positively related to the rate of learning. The a_3 parameter reflects the rate of initial learning (i.e., the learning from one trial). This is because during the first trial, $t_n = 1$ which leaves the above equation as a function of a_3. The c parameter is the Weibull shape parameter, and it reflects the hazard rate. If c is less than 1.0, then there is monotonically decreasing hazard; whereas if c is greater than 1.0, then there is monotonically increasing hazard with trials. And if $c = 1$, then the Weibull is equal to the Bower model and predicts constant hazard.

From the above equation, we have the following:

$$1 - \theta_{3C} = e^{-a_3 t_n^c}$$

$$\ln(1 - \theta_{3C}) = -a_3 t_n^c$$

$$\ln(\frac{1}{1 - \theta_{3C}}) = a_3 t_n^c$$

$$\ln\ln(\frac{1}{1 - \theta_{3C}}) = \ln a_3 + c\ln(t_n)$$

Let $y = \ln\ln\frac{1}{1-\theta_{3C}}$ and $x = \ln(t_n)$. We regress to fit A slope and B intercept, such that $\hat{c}$ corresponds to the slope (A) and $\hat{a}_3$ corresponds to e^B. Using this method, we can compare the Weibull fit vs. the observed parameter values for θ_{3C}

Based on the results displayed in Table 8, the Weibull fits closely resemble their respective observed parameters for both the Related and Unrelated condition. Furthermore, we can see yet again that subjects had an easier time learning Related triplets, as the learning rate was higher ($\hat{a}_3$ = 0.271) than that of the Unrelated condition ($\hat{a}_3$ = 0.102).

Let us now see how we can use the same model fitting method for θ_N. For this parameter, however, our coefficients of interest are d and f. We begin with the following expression and solve to find the slope and y-intercept of the model.

$$1 - \theta_N = 1 - e^{-dt_n^f}$$

$$\theta_N = e^{-dt_n^f}$$

In the equation above, the d parameter represents initial encoding, while f represents the shape parameter, which corresponds to the improvement rate in learning across trials. As in the previous equation for θ_{3C}, the term t_n represents the trial number. Hence, during the first trial, the t_n term disappears, and θ_N becomes a function of d. The larger the value for d, the smaller the value for θ_N, which indicated better encoding. As for f, it follows the same rules as c in the equation for θ_{3C}: when $f = 1$, hazard is constant; when $f > 1$, we have increasing hazard; and finally, when $f < 1$ we have decreasing hazard.

$$\ln\theta_N = -dt_n^f$$

$$\ln\frac{1}{\theta_N} = dt_n^f$$

$$\ln\ln\frac{1}{\theta_N} = \ln d + f\ln t_n$$

From the last expression above, we are able to find $y = AX + B$ and thus obtain the following.

$$B = \ln d$$

$$\hat{d} = e^B$$

$$\hat{f} = A$$

Following these steps, we obtain the predicted θ_N values as illustrated in Table 9. Again, the Weibull fits are very close to the observed values of θ_N. In regards to the learning rate, ($\hat{d}$), we find it is yet again higher for the Related condition (0.791) than that of the Unrelated condition (0.381), which aligns with what we have found thus far. In addition to this information, we also included the Weibull model's values of R^2 for Tables 8 and 9, which tell us how well the model is doing. We used the following formula to obtain the R^2 values.

$$R^2 = 1 - \frac{\Sigma(error)^2}{(n-1)VAR(\theta_{3C})}$$

The closer the R^2 is to 1.0, the better the model is performing and the smaller the model's associated error. The values we obtained for both the Related (0.907) and Unrelated (0.869) condition in Table 8 and Table 9 (0.915 and 0.976, respectively) show that the Weibull model is performing very well.

Table 8 Weibull Fits for θ_{3C} for Unrelated vs. Related Triplets

	Unrelated		Related	
Trial	θ_{3C}	$\hat{\theta}_{3C(Weibull)}$	θ_{3C}	$\hat{\theta}_{3C(Weibull)}$
1	0.107	0.097	0.270	0.237
2	0.148	0.172	0.256	0.349
3	0.281	0.236	0.500	0.430
4	0.223	0.293	0.475	0.494
5	0.366	0.344	0.550	0.547
6	0.433	0.391	0.639	0.591
7	0.363	0.433	0.623	0.628
8	0.543	0.472	0.640	0.661
R^2	0.869		0.907	
$\hat{a}_3$	0.102		0.271	
$\hat{c}$	0.882		0.666	

Table 9 Weibull Fits for θ_N for Unrelated vs. Related Triplets

	Unrelated		Related	
Trial	θ_N	$\hat{\theta}_{N(Weibull)}$	θ_N	$\hat{\theta}_{N(Weibull)}$
1	0.684	0.683	0.426	0.453
2	0.536	0.545	0.401	0.332
3	0.460	0.450	0.243	0.262
4	0.365	0.380	0.173	0.215
5	0.344	0.324	0.197	0.181
6	0.267	0.280	0.176	0.155
7	0.291	0.244	0.129	0.134
8	0.177	0.213	0.110	0.117
R^2	0.976		0.915	
$\hat{d}$	0.381		0.791	
$\hat{f}$	0.674		0.479	

Now that we've obtained the Weibull fits for both θ_{3C} and θ_N, we can subtract both quantities from 1 and obtain the partial learning parameter, θ_P. Table 10 displays all three observed and predicted parameters for both Unrelated and Related triplets. Here, we notice two things: firstly, the Weibull fits for θ_P are very close to their observed counterparts; secondly, the values of θ_P and $\hat{\theta}_P$ change differently across trials between the two groups. For the Unrelated condition, θ_P and $\hat{\theta}_P$ increase, while for the Related condition, θ_P and $\hat{\theta}_P$ decrease. This further speaks to the subjects' ease with learning the Related triplets over the Unrelated ones.

Table 10 Weibull fits for Complete, Partial, and Nonstorage of Order Learning

	Unrelated						Related					
Trial	θ_{3C}	$\hat{\theta_{3C}}$	θ_P	$\hat{\theta_P}$	θ_N	$\hat{\theta_N}$	θ_{3C}	$\hat{\theta_{3C}}$	θ_P	$\hat{\theta_P}$	θ_N	$\hat{\theta_N}$
1	0.107	0.097	0.209	0.220	0.684	0.683	0.270	0.237	0.304	0.310	0.426	0.453
2	0.148	0.172	0.316	0.283	0.536	0.545	0.256	0.349	0.343	0.319	0.401	0.332
3	0.281	0.236	0.259	0.314	0.460	0.450	0.500	0.430	0.257	0.308	0.243	0.262
4	0.223	0.293	0.412	0.327	0.365	0.380	0.475	0.494	0.352	0.291	0.173	0.215
5	0.366	0.344	0.290	0.332	0.344	0.324	0.550	0.547	0.253	0.272	0.197	0.181
6	0.433	0.391	0.300	0.329	0.267	0.280	0.639	0.591	0.185	0.254	0.176	0.155
7	0.363	0.433	0.346	0.323	0.291	0.244	0.623	0.628	0.248	0.238	0.129	0.134
8	0.543	0.472	0.280	0.315	0.177	0.213	0.640	0.661	0.250	0.222	0.110	0.117

Section II (Participant Subgroups)

We divided our 81 subjects into three subgroups to help us see how participants performed compared to one another. Subjects were grouped based on the number of errors they made throughout the study: Group 1 consists of participants who made between 1 and 37 total errors; Group 2 consists of participants who made between 38 and 71 errors; and finally, Group 3 consists of participants who made between 72 and 151 errors. Table 11 below shows the pooled frequencies across AFC tasks for Unrelated and Related triplets over the course of the experiment's eight trials. We can see that for both conditions, as we move from Group 1 to Group 3, the values for n_1 decrease while those for n_4 increase.

Table 11 Pooled Frequencies for Order Learning of Unrelated and Related Items by Group.

	Unrelated											
	Group 1				Group 2				Group 3			
Trial	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4
1	49	23	13	23	25	32	13	38	24	13	17	54
2	62	22	16	8	35	34	13	26	24	13	19	52
3	76	20	5	7	43	24	13	28	38	18	10	42
4	76	25	5	2	47	26	16	19	24	32	15	37
5	90	10	4	4	58	32	8	10	32	21	13	42
6	100	10	4	4	66	23	5	14	35	25	17	31
7	91	15	2	0	63	22	8	15	32	27	16	33
8	102	6	0	0	82	14	6	6	45	25	14	24
	Related											
	Group 1				Group 2				Group 3			
Trial	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4
1	69	17	13	9	44	19	18	27	36	28	14	30
2	72	24	6	6	45	22	16	25	35	23	17	33
3	94	9	2	3	76	13	9	10	46	25	9	28
4	93	11	2	2	73	19	8	8	46	27	17	18
5	99	7	1	1	79	18	2	9	50	26	8	24
6	101	5	1	1	84	11	6	7	63	20	3	22
7	105	2	1	1	83	20	1	4	58	22	12	16
8	102	7	0	0	92	11	0	5	64	24	7	13

As we did in the previous section, we now use the pooled frequencies from Table 11 to obtain the EO model parameters for each group and condition, as displayed in Table 12. For both conditions, there are stark differences across the three groups when we look at how much order knowledge they acquire by the time they finish the memory task. For the Unrelated condition, Group 1 reaches a θ_{3C} value of 0.857, followed by Group 2 (0.605), and Group 3 (0.177). As for the Related condition, Group 1 reaches a θ_{3C} value of 0.840, followed by Group 2 (0.724), and Group 3 (0.357). On average, Groups 2 and 3 performed better with the Related items, while Group 1 had an excellent performance with both the Related and Unrelated items.

Table 12 EO Parameters for Unrelated and Related Items by Group.

	Unrelated											
	Group 1				Group 2				Group 3			
Trial	θ_{3C}	θ_{2C}	θ_{1C}	θ_N	θ_{3C}	θ_{2C}	θ_{1C}	θ_N	θ_{3C}	θ_{2C}	θ_{1C}	θ_N
1	0.229	0.184	0.165	0.422	0.037	0.205	0.101	0.657	0.060	0.038	0.050	0.852
	(0.072)	(0.095)	(0.091)	(0.073)	(0.031)	(0.083)	(0.075)	(0.075)	(0.039)	(0.033)	(0.045)	(0.058)
2	0.348	0.142	0.349	0.161	0.063	0.311	0.155	0.471	0.059	0.037	0.060	0.844
	(0.074)	(0.088)	(0.091)	(0.051)	(0.047)	(0.097)	(0.095)	(0.075)	(0.039)	(0.033)	(0.049)	(0.061)
3	0.500	0.264	0.099	0.138	0.169	0.190	0.136	0.504	0.167	0.101	0.056	0.676
	(0.075)	(0.088)	(0.063)	(0.045)	(0.068)	(0.094)	(0.086)	(0.075)	(0.063)	(0.069)	(0.049)	(0.069)
4	0.456	0.357	0.134	0.053	0.185	0.193	0.266	0.357	0.034	0.183	0.128	0.655
	(0.079)	(0.094)	(0.065)	(0.030)	(0.072)	(0.101)	(0.104)	(0.072)	(0.029)	(0.083)	(0.086)	(0.078)
5	0.712	0.113	0.088	0.086	0.233	0.426	0.148	0.194	0.099	0.114	0.079	0.708
	(0.059)	(0.062)	(0.052)	(0.036)	(0.081)	(0.107)	(0.080)	(0.054)	(0.054)	(0.072)	(0.063)	(0.073)
6	0.736	0.104	0.081	0.079	0.384	0.300	0.070	0.246	0.097	0.150	0.187	0.565
	(0.055)	(0.057)	(0.048)	(0.034)	(0.077)	(0.092)	(0.055)	(0.055)	(0.056)	(0.086)	(0.099)	(0.080)
7	0.679	0.232	0.072	0.017	0.365	0.243	0.115	0.277	0.072	0.170	0.162	0.595
	(0.066)	(0.074)	(0.046)	(0.017)	(0.076)	(0.095)	(0.074)	(0.061)	(0.049)	(0.088)	(0.095)	(0.080)
8	0.857	0.104	0.026	0.013	0.605	0.149	0.123	0.123	0.177	0.201	0.181	0.441
	(0.047)	(0.048)	(0.025)	(0.013)	(0.067)	(0.074)	(0.065)	(0.044)	(0.070)	(0.099)	(0.097)	(0.075)
	Related											
	Group 1				Group 2				Group 3			
Trial	θ_{3C}	θ_{2C}	θ_{1C}	θ_N	θ_{3C}	θ_{2C}	θ_{1C}	θ_N	θ_{3C}	θ_{2C}	θ_{1C}	θ_N
1	0.454	0.111	0.257	0.178	0.204	0.094	0.206	0.496	0.091	0.222	0.147	0.540
	(0.071)	(0.075)	(0.083)	(0.054)	(0.067)	(0.070)	(0.094)	(0.079)	(0.056)	(0.095)	(0.092)	(0.077)
2	0.429	0.319	0.129	0.124	0.197	0.139	0.203	0.461	0.107	0.128	0.167	0.598
	(0.078)	(0.095)	(0.070)	(0.044)	(0.069)	(0.086)	(0.097)	(0.077)	(0.058)	(0.080)	(0.093)	(0.081)
3	0.758	0.123	0.055	0.065	0.555	0.101	0.150	0.194	0.188	0.249	0.083	0.480
	(0.056)	(0.059)	(0.041)	(0.030)	(0.066)	(0.066)	(0.071)	(0.055)	(0.072)	(0.096)	(0.065)	(0.070)
4	0.732	0.158	0.061	0.050	0.481	0.199	0.159	0.160	0.168	0.192	0.301	0.339
	(0.060)	(0.065)	(0.043)	(0.027)	(0.073)	(0.088)	(0.077)	(0.050)	(0.071)	(0.102)	(0.106)	(0.070)
5	0.820	0.106	0.043	0.031	0.545	0.259	0.044	0.152	0.215	0.290	0.082	0.414
	(0.051)	(0.052)	(0.033)	(0.021)	(0.072)	(0.079)	(0.039)	(0.043)	(0.075)	(0.098)	(0.064)	(0.067)
6	0.855	0.074	0.040	0.031	0.647	0.104	0.109	0.140	0.384	0.233	0.042	0.341
	(0.045)	(0.044)	(0.030)	(0.021)	(0.062)	(0.063)	(0.060)	(0.046)	(0.073)	(0.084)	(0.038)	(0.060)
7	0.906	0.035	0.030	0.029	0.564	0.323	0.038	0.074	0.319	0.187	0.192	0.302
	(0.036)	(0.028)	(0.023)	(0.019)	(0.074)	(0.079)	(0.034)	(0.031)	(0.074)	(0.095)	(0.091)	(0.066)
8	0.840	0.121	0.026	0.013	0.724	0.169	0.027	0.080	0.357	0.294	0.107	0.241
	(0.049)	(0.051)	(0.025)	(0.013)	(0.060)	(0.062)	(0.027)	(0.032)	(0.077)	(0.096)	(0.071)	(0.057)

Now that we've looked at how the three groups differ in terms of order knowledge, we follow the same process for item storage. Table 13 illustrates the pooled frequencies for item storage across groups. At first glance, we do not see as stark of a difference in the AFC responses for the three groups. The values of n_1 for all three groups are close to each other by the time they finish the memory task, with Group 1 reaching 107, Group 2 with 104, and Group 3 with 100. Complementing this information, n_4 is very small for all groups by trial 8, with Group 1 reaching 0, Group 2 reaching 1, and Group 3 reaching 2.

Table 13 Pooled Frequencies for Item Storage.

	Group 1				Group 2				Group 3			
Trial	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4	n_1	n_2	n_3	n_4
1	98	7	3	0	75	17	10	6	75	11	12	10
2	95	10	2	1	82	14	8	4	72	20	12	4
3	106	1	1	0	102	1	3	2	84	13	8	3
4	108	0	0	0	102	2	3	1	89	8	8	3
5	107	1	0	0	97	6	4	1	82	15	9	2
6	107	1	0	0	102	3	1	2	89	11	6	2
7	108	0	0	0	104	3	1	0	98	6	4	0
8	107	1	0	0	104	3	0	1	100	5	1	2

We can get a clearer picture of the groups' performance by looking at their item storage parameters as shown in Table 14. All three groups had a better performance while learning item content information as opposed to item order. By the end of the experiment, Group 1's θ_{Si} was 0.941, Group 2's was 0.805, and Group 3's was 0.593. These item storage values are higher than those of order learning (θ_{3C}) for all three groups. Furthermore, the values of θ_{Ni} are all very close to zero (0.014, 0.060, and 0.104).

Table 14 IS Model Parameters by Group.

	Group 1			Group 2			Group 3		
Trial	θ_{Si}	θ_{Fi}	θ_{Ni}	θ_{Si}	θ_{Fi}	θ_{Ni}	θ_{Si}	θ_{Fi}	θ_{Ni}
1	0.809 (0.051)	0.161 (0.054)	0.030 (0.028)	0.512 (0.071)	0.262 (0.090)	0.225 (0.073)	0.542 (0.065)	0.140 (0.072)	0.318 (0.071)
2	0.758 (0.058)	0.190 (0.061)	0.052 (0.033)	0.602 (0.067)	0.237 (0.081)	0.160 (0.063)	0.460 (0.073)	0.367 (0.096)	0.174 (0.073)
3	0.929 (0.032)	0.052 (0.030)	0.019 (0.017)	0.881 (0.038)	0.056 (0.034)	0.063 (0.031)	0.628 (0.064)	0.240 (0.078)	0.132 (0.059)
4	0.957 (0.025)	0.031 (0.022)	0.012 (0.012)	0.878 (0.039)	0.073 (0.039)	0.048 (0.029)	0.707 (0.058)	0.165 (0.067)	0.128 (0.056)
5	0.941 (0.029)	0.045 (0.027)	0.014 (0.014)	0.805 (0.051)	0.136 (0.055)	0.060 (0.037)	0.593 (0.068)	0.303 (0.082)	0.104 (0.056)
6	0.942 (0.029)	0.045 (0.027)	0.014 (0.013)	0.877 (0.040)	0.067 (0.037)	0.056 (0.029)	0.692 (0.061)	0.212 (0.070)	0.096 (0.049)
7	0.957 (0.025)	0.031 (0.022)	0.012 (0.012)	0.897 (0.038)	0.081 (0.038)	0.022 (0.020)	0.814 (0.049)	0.153 (0.053)	0.032 (0.030)
8	0.941 (0.029)	0.045 (0.027)	0.014 (0.014)	0.805 (0.051)	0.136 (0.055)	0.060 (0.037)	0.593 (0.068)	0.303 (0.082)	0.104 (0.056)

Although we previously obtained the β_N values for all subjects (Table 7), we decided to look at how these values differ across groups as well. According to Table 15, the β_N values all decrease in a non-linear fashion. In other words, we are seeing the same trend that we observed when we looked at all subjects in Table 7. From this information, we can determine that sticking with the Weibull model is the best path of action for our subgroup-related analyses.

Table 15 β_N Values by Group.

Trial	$\beta_{N(G1)}$	$\beta_{N(G2)}$	$\beta_{N(G3)}$
1	0.700	0.404	0.252
2	0.631	0.313	0.126
3	0.539	0.293	0.158
4	0.539	0.289	0.156
5	0.443	0.293	0.104
6	0.492	0.238	0.117
7	0.490	0.220	0.107
8	0.464	0.246	0.125

In the same fashion as for Table 7, we obtained Kendall τ correlation coefficients for the relationship between the β values and the trial number. We obtained a coefficient for each group. We obtained a Kendall τ value of: -0.7638 for Group 1 (with a probability of 0.9999863), -0.7638 for Group 2 (with a probability of 0.9999863), and -0.5 for Group 3 (with a probability of 0.995935). We are seeing a pattern here that is reminiscent of what we previously: for all three groups, the relationship between the trial number and their corresponding β values is negative and of moderate strength (though less so for Group 3).

We obtained the EO model parameters' Weibull fits for each group and for both Unrelated and Related triplets. These results are displayed in Table 16, where we also included the predicted Weibull fit for the partial learning parameter, θ_P. In addition, we included the R^2 values for the $\hat{\theta_{3C}}$ and $\hat{\theta_N}$ models, along with their respective slopes and learning rates. For both Unrelated and Related triads, the we find the learning rate to be higher for the Weibull fits for θ_N. Generally, we also see that the learning rate tends to be higher for the Related triads, again supporting the idea that subjects had an easier time learning in the related condition. Lastly, the R^2 values are very high (with the exception of 0.113), which shows that the Weibull model is performing well.

Table 16 Weibull Model Fits for Order Learning by Group

	Unrelated								
	Group 1			Group 2			Group 3		
Trial	$\hat{\theta_{3C}}$	$\hat{\theta_P}$	$\hat{\theta_N}$	$\hat{\theta_{3C}}$	$\hat{\theta_P}$	$\hat{\theta_N}$	$\hat{\theta_{3C}}$	$\hat{\theta_P}$	$\hat{\theta_N}$
1	0.214	0.399	0.387	0.030	0.307	0.663	0.056	0.073	0.871
2	0.363	0.421	0.216	0.082	0.405	0.513	0.070	0.138	0.792
3	0.479	0.390	0.131	0.144	0.445	0.411	0.079	0.191	0.730
4	0.572	0.344	0.084	0.212	0.451	0.337	0.086	0.237	0.677
5	0.646	0.298	0.056	0.282	0.437	0.281	0.092	0.278	0.630
6	0.707	0.255	0.038	0.351	0.413	0.236	0.097	0.314	0.589
7	0.756	0.218	0.026	0.419	0.381	0.200	0.102	0.346	0.552
8	0.797	0.185	0.018	0.484	0.345	0.171	0.106	0.375	0.519
R^2	0.911	-	0.935	0.908	-	0.882	0.113	-	0.844
Learning Rate	0.910	-	0.691	1.473	-	0.701	0.316	-	0.748
Encoding	0.240	-	0.950	0.031	-	0.411	0.058	-	0.139
	Related								
	Group 1			Group 2			Group 3		
Trial	$\hat{\theta_{3C}}$	$\hat{\theta_P}$	$\hat{\theta_N}$	$\hat{\theta_{3C}}$	$\hat{\theta_P}$	$\hat{\theta_N}$	$\hat{\theta_{3C}}$	$\hat{\theta_P}$	$\hat{\theta_N}$
1	0.407	0.404	0.189	0.182	0.294	0.524	0.076	0.318	0.606
2	0.571	0.323	0.106	0.305	0.339	0.356	0.130	0.362	0.508
3	0.675	0.255	0.070	0.402	0.341	0.257	0.175	0.380	0.445
4	0.747	0.204	0.049	0.482	0.326	0.192	0.216	0.385	0.399
5	0.799	0.164	0.037	0.549	0.304	0.147	0.252	0.384	0.364
6	0.839	0.133	0.028	0.606	0.280	0.114	0.286	0.380	0.334
7	0.869	0.109	0.022	0.655	0.255	0.090	0.317	0.373	0.310
8	0.893	0.089	0.018	0.696	0.233	0.071	0.346	0.365	0.289
R^2	0.853	-	0.976	0.945	-	0.906	0.838	-	0.792
Learning Rate	0.699	-	0.426	0.857	-	0.677	0.807	-	0.437
Encoding	0.522	-	1.668	0.201	-	0.646	0.079	-	0.501

4.2 Theoretical Framework for Order Learning

One important aspect of our findings is how the Weibull model parameters behave across the conditions and what that means. For θ_{3C}, the a_3 parameter is a measure of storage in the $3C$ state after 1 training trial because $\theta_{3C(trial1)} = 1 - \exp(-a_3)$. So, the bigger the value for a_3, the bigger the value for θ_{3C}. For example, if a_3 = 0.2, then θ_{3C} would be 0.181; but, if a_3 = 0.1, then θ_{3C} would be 0.095. The c parameter is related to the level of learning hazard, meaning that the rate of improvement in order knowledge increases from trial 1 at a rate that is positively correlated with c. Note that in Table 8, for the pooling over all participants, $\theta_{3C(trial1)}$ is larger for the Related condition in comparison to the Unrelated condition, and this difference is connected to the a_3 parameter being larger for the Related triplets. Additionally, note that the c value for both conditions is less than 1, showing that the improvement in order learning is slowing down. If the learning rate were a constant for unrelated triplets and if the value for a_3 were the same, then the predicted value for θ_{3C} after the eight trial should have been 0.558, which is too large. Consequently, for the Unrelated condition the value for c is slightly reduced from 1 to the value of 0.882. However, for the Related triplets, the predicted $\theta_{3C(trial8)}$ based on the a_3 value and constant learning rate would result in predicting a value of 0.886, which is much too large. Consequently, the c value is reduced from 1 to a greater degree (i.e., it is 0.666).

The quantity $1 - \theta_N$ is a measure of some order knowledge. Based on the item storage (IS) model, the quantity $\theta_{Si}^3 + 3(1 - \theta_{Si}) * \theta_{Si}^2$ is a measure of at least two items of a triplet being stored. If the first quantity is less than the second, then there is statistical evidence that order knowledge is less than what would be expected by of item storage alone (i.e., to learn order it requires more than the two of at least two items). We can assess if the random samples from the posterior distribution of $1 - \theta_N$ is less than a sample from the distribution for $\theta_{Si}^3 + 3 * (1 - \theta_{Si}) * \theta_{Si}^2$. From the Monte Carlo samples for Group 1 for trials 1 and 2 combine the Bayes factor $BF_{10} = 55$, which shows that order knowledge is less than what would be expected from only item storage. The Bayes factor from for Group 1 for trials 3 and 4, trials 5 and 6, and trials 7 and 8 were respectively 7,806, 1,764, and 1.41. So, for Group 1 there was less order knowledge than item knowledge for the first six trials, and then the high performing group was approaching mastery of both item and order information. Similar analyses were also performed for Groups 2 and 3. For Group 2 the four consecutive Bayes factors were 11,490, 5,317, 1,764, and 23,323; whereas for Group 3 the four consecutive Bayes factors were 11,490, 38,684, 1,421, and 39,183. Thus for both the middle and lowest performing groups there is clearly more difficulty learning order than in learning at least two items of triplet.

The Weibull model for θ_N is described by the following expression: $\theta_N = \exp(-d * t_n^f)$. It is a similar model, but now the $\theta_{N(trial1)}$ is $\exp(-d)$, and it has an inverse effect. That is, if d is larger then θ_N is smaller (i.e., $1 - \theta_N$ is larger). So, learning here corresponds to the decrease in θ_N. The f parameter has to do with the rate of this reduction throughout the eight trials. A constant rate would be equivalent to f=1. As with θ_{3C}, note that θ_N decreases at a slower rate over time. From Table 9 we see that the d parameter is larger for the Related condition, and that corresponds to the fact that θ_N is smaller in that condition in comparison to the Unrelated condition. Yet, the rate of the decline in θ_N slows down for both condition which is reflect by the fact that f is less than 1 for both conditions. Moreover, the slowing is more pronounced for the Related condition. Finally, in Table 16 the story described above holds for groups 1 and 2. However, for group 3 the performance of the participants is nearly equally low after trial 1 for the θ_{3C} parameter. Despite this, the rate of the subsequent order learning is not slower for the Related triplets. Nonetheless, the above story for the reduction in θ_N for group 3 is similar to the overall grouping of participants.

4.3 General Discussion

The goal of this thesis was to explore the measurement of memory for order information, which we were able to do through the four experiments we conducted. The main feature of our approach was to use a modeling of the states of order knowledge that could be exhibited when the participant is probed about a triplet of items. The resulting EO model identified four states of order knowledge: complete storage, which occurs with probability θ_{3C}; storage of two correct order relationships, which occurs with probability θ_{2C}; storage of a single order relationship, which occurs with probability θ_{1C}; and finally, the non-storage of any order information which, occurs with probability θ_N. The EO model successfully measured these four information states in all four experiments, and it gave us a clearer picture of how order knowledge changes over time.

As for Experiments 3 and 4, the performance on the EO task and model were compared to the Chechile-Soraci model for measuring the storage of single items. We found evidence from both of these studies that the learning of order is more than simply learning the individual events. If the three members of an arbitrary triplet are not stored, then there cannot be any learning of order. In general, learning items is relatively easy compared to the learning of their order. Our research also shows that each state of order knowledge occurs. While it is clear that the no knowledge state and complete-order require, respectively, zero and three order relationships stored, the question is: how can 1-correct state and 2-correct state be achieved? The 2-correct state requires all three items to be stored and for two of three items to be in a known order-relationship. For the 1-correct state can occur when two items are stored and the order relationship between those items is also stored. For the 1-correct state, the third item might or might not be stored.

Furthermore, our experiments show that the spacing among the items affect order learning. Generally, adjacent items are easier to learn than far items. However, order learning is better in a far spacing condition in comparison to a near spacing condition. It is also established that it is easier to learn the order of items that are semantically similar rather dissimilar. We've seen throughout all four experiments that participants have had a superior performance with Related triplets than with Unrelated ones. Finally, learning of order is difficult and was not perfect for most participants, even after eight training trials. On the other hand, we saw in Experiments 3 and 4 that participants performed substantially better on item-content questions. Comparing the EO model to the IS model in Experiment 3 allowed us to see that memory for order and memory for content are different. In Experiment 4, we gathered more support for this idea by looking at the differences in learning rates and parameters across subgroups.

Yet, there are wide individual differences for the learning of order. The top one third of the participants could perfectly order a triplet of unrelated items 80% of the time; whereas the bottom one third of the participants could only correctly order a triplet on 10% of the occasions after eight training trails. It is possible that these differences in individual performance could be due to some students using mnemonics or the method of loci to improve their results. Another important finding is the occurrence of the θ_{2C} state. Across all four experiments, we were able to obtain a full probability distribution for each of the four states of order knowledge. In particular, us finding evidence for the existence of θ_{2C} is problematic for both the chaining model and the position model. Not, however, for the perturbation model, which could account for our results (Lee and Estes, 1981; Lee, 1992).

Lastly, we showed that the learning rate for order is not constant over time. We found this due to our use of a Weibull learning model to study order learning. The Weibull model had a two parameters: one was measuring the degree of learning after a single trial, and the other parameter was tapping the rate of subsequent order learning. In general, factors such item relatedness affect the initial trial 1 learning. The subsequent learning of order slows down.

Chapter 5

Appendix

5.0.1 Lists

Experiment 1: https://doi.org/10.7910/DVN/RSK1UT

Experiment 2: https://doi.org/10.7910/DVN/5ELTXC

Experiment 3: https://doi.org/10.7910/DVN/4A63FO

Experiment 4: https://doi.org/10.7910/DVN/TUUFSF

5.0.2 Order Learning Triplets: Unrelated

Experiment 1: https://doi.org/10.7910/DVN/RSK1UT

Experiment 2: https://doi.org/10.7910/DVN/5ELTXC

Experiment 3: https://doi.org/10.7910/DVN/4A63FO

Experiment 4: https://doi.org/10.7910/DVN/TUUFSF

5.0.3 Order Learning Triplets: Related

Experiment 1: https://doi.org/10.7910/DVN/RSK1UT

Experiment 2: https://doi.org/10.7910/DVN/5ELTXC

Experiment 3: https://doi.org/10.7910/DVN/4A63FO

Experiment 4: https://doi.org/10.7910/DVN/TUUFSF

5.0.4 Item Learning Stimuli

Experiment 3: https://doi.org/10.7910/DVN/4A63FO

Experiment 4: https://doi.org/10.7910/DVN/TUUFSF

5.0.5 Special Cases with the MLE

Case 1

In the case where $n_1 = 0$, $n_2 = 0$, $n_3 = 0$, and $n_4 > 0$:

$\widehat{\theta_N} = 1$

$\widehat{\theta_{1C}} = 0$

$\widehat{\theta_{2C}} = 0$

$\widehat{\theta_{3C}} = 0$

Case 2

In the case where $n_1 = 0$, $n_2 = 0$, $n_3 > 0$, and $n_4 = 0$:

$\widehat{\theta_N} = 0$

$\widehat{\theta_{1C}} = 1$

$\widehat{\theta_{2C}} = 0$

$\widehat{\theta_{3C}} = 0$

Case 3

In the case where $n_1 = 0$, $n_2 > 0$, $n_3 = 0$, and $n_4 = 0$:

$\widehat{\theta_N} = 0$

$\widehat{\theta_{1C}} = 0$

$\widehat{\theta_{2C}} = 1$

$\widehat{\theta_{3C}} = 0$

Case 4

In the case where $n_1 > 0$, $n_2 = 0$, $n_3 = 0$, and $n_4 = 0$:

$\widehat{\theta_N} = 0$

$\widehat{\theta_{1C}} = 0$

$\widehat{\theta_{2C}} = 0$

$\widehat{\theta_{3C}} = 1$

Case 5

In the case where $n_1 > 0$, $n_2 > 0$, $n_3 = 0$, and $n_4 = 0$:

$\widehat{\theta_N} = 0$

$\widehat{\theta_{1C}} = 0$

$$\widehat{\theta_{3C}} = \begin{cases} 0, & \text{if } n_2 \geq n_1 \& \ \widehat{\theta_{2C}} = 1 \\ \frac{n_1 - n_2}{n_1 + n_2} & \text{if } n_2 < n_1 \& \ \widehat{\theta_{2C}} = \frac{2n_2}{n_1 + n_2} \end{cases}$$

Case 6

In the case where $n_1 > 0$, $n_2 = 0$, $n_3 > 0$, and $n_4 = 0$:

$\widehat{\theta_N} = 0$

$\widehat{\theta_{2C}} = 0$

$$(\widehat{\theta_{3C}}, \widehat{\theta_{1C}}) = \begin{cases} (\frac{2n_1 - n_3}{2n_1 + 2n_3}, \frac{3n_3}{2n_1 + 2n_3}), & \text{if } 2n_1 \geq n_3 \\ (0, 1), & \text{if } 2n_1 < n_3 \end{cases}$$

Case 7

In the case where $n_1 > 0$, $n_2 = 0$, $n_3 = 0$, and $n_4 > 0$:

$\widehat{\theta_{2C}} = 0$

$\widehat{\theta_{1C}} = 0$

$$(\widehat{\theta_{3C}}, \widehat{\theta_N}) = \begin{cases} (\frac{n_1 - n_4}{n_1 + n_4}, \frac{2n_4}{n_1 + n_4}), & \text{if } n_1 \geq n_4 \\ (0, 1), & \text{if } n_1 < n_4 \end{cases}$$

Case 8

In the case where $n_1 = 0$, $n_2 > 0$, $n_3 > 0$, and $n_4 = 0$:

$\widehat{\theta_{3C}} = 0$

$\widehat{\theta_N} = 0$

$$(\widehat{\theta_{2C}}, \widehat{\theta_{1C}}) = \begin{cases} (\frac{n_2-2n_3}{n_2+n_3}, \frac{3n_3}{n_2+n_3}), & \text{if } n_2 \geq 2n_3 \\ (0, 1), & \text{if } n_2 < 2n_3 \end{cases}$$

Case 9

In the case where $n_1 = 0$, $n_2 > 0$, $n_3 = 0$, and $n_4 > 0$:

$\widehat{\theta_{3C}} = 0$

$\widehat{\theta_{1C}} = 0$

$$(\widehat{\theta_{2C}}, \widehat{\theta_N}) = \begin{cases} (\frac{n_2-n_4}{n_4+n_2}, \frac{2n_4}{n_4+n_2}), & \text{if } n_2 \geq n_4 \\ (0, 1), & \text{if } n_2 < n_4 \end{cases}$$

Case 10

In the case where $n_1 = 0$, $n_2 = 0$, $n_3 > 0$, and $n_4 > 0$:

$\widehat{\theta_{3C}} = 0$

$\widehat{\theta_{2C}} = 0$

$$(\widehat{\theta_{1C}}, \widehat{\theta_N}) = \begin{cases} (\frac{n_3-n_4}{n_3+n_4}, \frac{2n_4}{n_3+n_4}), & \text{if } n_3 \geq n_4 \\ (0, 1), & \text{if } n_3 < n_4 \end{cases}$$

Case 11

In the case where $n_1 > 0$, $n_2 > 0$, $n_3 > 0$, and $n_4 = 0$:

$\widehat{\theta_N} = 0$

$$(\widehat{\theta_{1C}}, \widehat{\theta_{2C}}, \widehat{\theta_{3C}}) = \begin{cases} (\frac{3n_3}{n_1+n_2+n_3}, \frac{2n_2-2n_3}{n_1+n_2+n_3}, \frac{n_1-n_2}{n_1+n_2+n_3}), & \text{if } n_1 \geq n_2 \geq n_3 \\ (1,0,0), & \text{if } n_1 < n_2, n_2 < n_3 \\ (\frac{3n_3-n_3+n_1}{n_1+n_2+n_3}, \frac{2n_2-2n_3}{n_1+n_2+n_3}, 0) & \text{if } n_1 < n_2, n_2 > n_3 \\ (\frac{2n_2+n_3}{n_1+n_2+n_3}, 0, \frac{n_1-n_2}{n_1+n_2+n_3}) & \text{if } n_1 \geq n_2, n_2 < n_3 \end{cases}$$

Case 12

In the case where $n_1 > 0$, $n_2 > 0$, $n_3 = 0$, and $n_4 > 0$:

$\widehat{\theta_{1C}} = 0$

$$(\widehat{\theta_N}, \widehat{\theta_{2C}}, \widehat{\theta_{3C}}) = \begin{cases} (\frac{2n_4}{n_1+n_2+n_4}, \frac{2n_2-\frac{2}{3}n_4}{n_1+n_2+n_4}, \frac{n_1-n_2-\frac{1}{3}n_4}{n_1+n_2+n_4}), & \text{if } n_1+n_2 > n_4, n_2 > \frac{1}{3}n_4, n_1 > n_2+\frac{1}{3}n_4 \\ (1,0,0), & \text{if } n_1+n_2 \leq n_4 \\ (\frac{2n_4}{n_1+n_2+n_4}, 0, \frac{n_1+n_2-n_4}{n_1+n_2+n_4}) & \text{if } n_1+n_2 > n_4, n_2 \leq \frac{1}{3}n_4 \\ (\frac{2n_4}{n_1+n_2+n_4}, \frac{n_1+n_2-n_4}{n_1+n_2+n_4}, 0) & \text{if } n_1+n_3 > n_4, n_2 > \frac{1}{3}n_4, n_1 \leq n_2+\frac{1}{3}n_4 \end{cases}$$

Case 13

In the case where $n_1 > 0$, $n_2 = 0$, $n_3 > 0$, and $n_4 > 0$:

$\widehat{\theta_{2C}} = 0$

$$(\widehat{\theta_N}, \widehat{\theta_{1C}}, \widehat{\theta_{3C}}) = \begin{cases} (\frac{2n_4}{n_1+n_3+n_4}, \frac{3n_3-n_4}{n_1+n_3+n_4}, \frac{n_1-2n_3}{n_1+n_3+n_4}), & \text{if } n_1+n_3 > n_4, n_3 > \frac{1}{3}n_4, n_1 > 2n_3 \\ (1,0,0), & \text{if } n_1+n_3 \leq n_4 \\ (\frac{2n_4}{n_1+n_3+n_4}, 0, \frac{n_1+n_3-n_4}{n_1+n_3+n_4}) & \text{if } n_1+n_3 > n_4, n_3 < \frac{1}{3}n_4 \\ (\frac{2n_4}{n_1+n_3+n_4}, \frac{n_1+n_3-n_4}{n_1+n_3+n_4}, 0) & \text{if } n_1+n_3 > n_4, n_3 \geq \frac{1}{3}n_4, n_1 \leq 2n_3 \end{cases}$$

Case 14

In the case where $n_1 = 0$, $n_2 > 0$, $n_3 > 0$, and $n_4 > 0$:

$\widehat{\theta_{3C}} = 0$

$$(\widehat{\theta_N}, \widehat{\theta_{1C}}, \widehat{\theta_{2C}}) = \begin{cases} (\frac{2n_4}{n_2+n_3+n_4}, \frac{3n_3-n_4}{n_2+n_3+n_4}, \frac{n_2-2n_3}{n_2+n_3+n_4}), & \text{if } n_2+n_3 > n_4, n_3 > \frac{1}{3}n_4, n_2 > 2n_3 \\ (1,0,0), & \text{if } n_2+n_3 \leq n_4 \\ (\frac{2n_4}{n_2+n_3+n_4}, 0, \frac{n_2+n_3-n_4}{n_2+n_3+n_4}) & \text{if } n_2+n_3 > n_4, n_3 < \frac{1}{3}n_4 \\ (\frac{2n_4}{n_2+n_3+n_4}, \frac{n_2+n_3-n_4}{n_2+n_3+n_4}, 0) & \text{if } n_2+n_3 > n_4, n_3 \geq \frac{1}{3}n_4, n_2 \leq 2n_3 \end{cases}$$

5.0.6 R Software Code for PPM: EO Model

The following R code implements 50,000 random samples for each of the four parameters of the EO model. The values of the pooled frequencies (n_1, n_2, n_3, n_4) have to be changed prior to running the program. After running the program, the output consists of vectors for each EO model parameter. The coherence probability of the model is also provided.

```
# The outcome cells frequencies must entered
# for each program run. These values for this case are:
n1=40
n2=30
n3=23
n4=6
theta3c=seq(1,50000,1)
theta2c=seq(1,50000,1)
theta1c=seq(1,50000,1)
thetaN=seq(1,50000,1)
a1=n1+1
a2=n2+1
```

```
a3=n3+1
a4=n4+1
aT=a1+a2+a3+a4
b1=rbeta(50000,a1,aT-a1)
b2=rbeta(50000,a2,aT-a1-a2)
b3=rbeta(50000,a3,a4)
phi1=b1
phi2=(1-b1)*b2
phi3=(1-b1)*(1-b2)*b3
phi4=(1-b1)*(1-b2)*(1-b3)
c=0
for (i in 1:50000) {
 if ((phi1[i]>=phi2[i]) & (phi2[i]>=phi3[i])
& (phi3[i]>=(1/3)*phi4[i]) & (phi4[i]<=.5)) {
  c=c+1
  thetaN[c]=2*phi4[i]
  theta1c[c]=3*(phi3[i])-phi4[i]
  theta2c[c]=2*(phi2[i]-phi3[i])
  theta3c[c]=phi1[i] - phi2[i]
   }
  else {c=c} }
theta3c=theta3c[1:c]
theta2c=theta2c[1:c]
theta1c=theta1c[1:c]
thetaN=thetaN[1:c]
probcoh=c/50000
```

5.0.7 R Software Code for PPM: IS Model

The software for implementing the item-storage analysis for the data about item knowledge in Experiment 3 is shown below.

```
# The model is a variation of the Chechile
# Soraci model. The n values need to be changed.
# for each run. These values for this run are:
n1=55
n2=25
n3=15
n4=5
alpha1=n1+1
alpha2=n2+1
alpha3=n3+1
alpha4=n4+1
alphaT=alpha1+alpha2+alpha3+alpha4
# The following create vectors for output
thetaSi=seq(1,50000,1)
thetaRi=seq(1,50000,1)
thetaFi=seq(1,50000,1)
thetaNi=seq(1,50000,1)
thetaf2=seq(1,50000,1)
# The random sample from Dirichlet are as follows
b1=rbeta(50000,alpha1, alphaT-alpha1)
b2=rbeta(50000,alpha2, alphaT-alpha1-alpha2)
b3=rbeta(50000,alpha3,alpha4)
phi1=b1
phi2=(1-b1)*b2
phi3=(1-b1)*(1-b2)*b3
phi4=(1-b1)*(1-b2)*(1-b3)
# The above phi1, phi2, phi3, and phi4 are random
# values from the posterior Dirichlet distribution.
# The following is the PPM code.
c=0
```

Experiment 1: Table of Pooled Frequencies

		Attempt			
		1st	2nd	3rd	4th
Probe Type	**AR1**	201	56	31	72
	AR2	270	38	21	31
	AU1	144	57	39	120
	AU2	154	70	48	88
	FR1	159	71	39	91
	FR2	218	60	28	54
	FU1	144	79	38	99
	FU2	201	70	28	61
	NR1	147	74	49	90
	NR2	203	57	30	70
	NU1	99	69	50	142
	NU2	130	75	40	115

Figure 5.1: Table of pooled frequencies by probe type and attempt number for Experiment 1.

```
for (i in 1:50000){
   if ((phi1[i]>=phi2[i])&(phi2[i]>=phi3[i])
   &(phi3[i]>=phi4[i]) &(phi4[i]<=.25)) {
   c=c+1
   thetaSi[c]=phi1[i]-phi2[i]
   thetaFi[c]=(2*phi2[i])+phi3[i]-(3*phi4[i])
   thetaNi[c]=4*phi4[i]
   thetaf2[c]=((6*(phi2[i]-phi4[i]))/thetaFi[c])-2
              }
   else {c=c}
}
thetaSi=thetaSi[1:c]
thetaFi=thetaFi[1:c]
thetaNi=thetaNi[1:c]
thetaf2=thetaf2[1:c]
probcoh=c/50000
```

AR: Test 1

n1 = 201	n2 = 56	n3 = 31	n4 = 72

P(coh)	0.8873
θ_3C	0.398475 (0.0389785)
θ_2C	0.1326491 (0.04887345)
θ_1C	0.07321901 (0.04415116)
θ_N	0.3956569 (0.03924879)

AR: Test 2

n1 = 270	n2 = 38	n3 = 21	n4 = 31

P(coh)	0.98028
θ_3C	0.6369706 (0.03471906)
θ_2C	0.09463636 (0.0406744)
θ_1C	0.09293272 (0.03962989)
θ_N	0.1754604 (0.02932614)

NR: Test 1

n1 = 147	n2 = 74	n3 = 49	n4 = 90

P(coh)	0.98534
θ_3C	0.1999225 (0.03959444)
θ_2C	0.1390564 (0.05849041)
θ_1C	0.1617646 (0.0614527)
θ_N	0.4992565 (0.04537057)

NR: Test 2

n1 = 203	n2 = 57	n3 = 30	n4 = 70

P(coh)	0.88102
θ_3C	0.4012998 (0.0390567)
θ_2C	0.1430605 (0.04912039)
θ_1C	0.07099431 (0.04337081)
θ_N	0.3846454 (0.03889083)

FR: Test 1

n1 = 159	n2 = 71	n3 = 39	n4 = 91

P(coh)	0.9101
θ_3C	0.2416306 (0.03983113)
θ_2C	0.1714186 (0.05533675)
θ_1C	0.08642316 (0.05061393)
θ_N	0.5005276 (0.04293848)

FR: Test 2

n1 = 218	n2 = 60	n3 = 28	n4 = 54

P(coh)	0.97168
θ_3C	0.4342983 (0.03975416)
θ_2C	0.1742687 (0.05065793)
θ_1C	0.09065598 (0.04578261)
θ_N	0.300777 (0.03654692)

Figure 5.2: Tables of Experiment 1's average parameter estimates for the EO model. The tables belong to the Adjacent, Near, and Far conditions for Related words. The standard deviations for each estimate are shown in parentheses.

AU: Test 1

n1 = 144	n2 = 57	n3 = 39	n4 = 120

P(coh)	0.4358
θ_3C	0.237366 (0.03658539)
θ_2C	0.08050184 (0.04349572)
θ_1C	0.04498988 (0.03471282)
θ_N	0.6371423 (0.04110054)

AU: Test 2

n1 = 154	n2 = 70	n3 = 48	n4 = 88

P(coh)	0.97468
θ_3C	0.2300109 (0.03907038)
θ_2C	0.124046 (0.05594644)
θ_1C	0.1572594 (0.05977373)
θ_N	0.4886837 (0.04495579)

NU: Test 1

n1 = 99	n2 = 69	n3 = 50	n4 = 142

P(coh)	0.60468
θ_3C	0.08133129 (0.0333605)
θ_2C	0.09200276 (0.04851194)
θ_1C	0.06159455 (0.04390105)
θ_N	0.7650714 (0.0440349)

NU: Test 2

n1 = 130	n2 = 75	n3 = 40	n4 = 115

P(coh)	0.61418
θ_3C	0.1511482 (0.03846436)
θ_2C	0.1736998 (0.05453574)
θ_1C	0.05621604 (0.04188547)
θ_N	0.618936 (0.04267727)

FU: Test 1

n1 = 144	n2 = 79	n3 = 38	n4 = 99

P(coh)	0.78874
θ_3C	0.1785733 (0.0401216)
θ_2C	0.2147944 (0.05622555)
θ_1C	0.06750834 (0.04557056)
θ_N	0.5391239 (0.04236359)

FU: Test 2

n1 = 201	n2 = 70	n3 = 28	n4 = 61

P(coh)	0.92612
θ_3C	0.3601012 (0.04083394)
θ_2C	0.2269862 (0.05241743)
θ_1C	0.07585383 (0.04388568)
θ_N	0.3370588 (0.03727487)

Figure 5.3: Tables of Experiment 1's average parameter estimates for the EO model. The tables belong to the Adjacent, Near, and Far conditions for Unrelated words. The standard deviations for each estimate are shown in parentheses.

Experiment 2: Table of Pooled Frequencies

		Attempt			
		1st	**2nd**	**3rd**	**4th**
Probe Type	**AR1**	163	72	25	91
	AR2	248	46	12	45
	AR3	260	39	17	35
	AR4	293	32	8	18
	AU1	123	59	42	127
	AU2	121	74	41	115
	AU3	214	55	21	61
	AU4	203	46	21	81
	FR1	102	87	49	113
	FR2	208	56	22	65
	FR3	198	67	30	56
	FR4	261	43	17	30
	FU1	136	97	28	90
	FU2	184	101	26	40
	FU3	205	80	17	49
	FU4	222	83	15	31
	NR1	126	72	42	111
	NR2	141	79	49	82
	NR3	229	42	33	47
	NR4	225	59	21	46
	NU1	103	60	52	136
	NU2	121	60	45	125
	NU3	156	69	36	90
	NU4	178	61	34	78

Figure 5.4: Table of pooled frequencies by probe type and attempt number for Experiment 2.

AR: Test 1

n1 = 163	n2 = 72	n3 = 25	n4 = 91

P(coh)	0.21624
θ_3C	0.2561514 (0.04092872)
θ_2C	0.2291923 (0.05121421)
θ_1C	0.03150531 (0.02733231)
θ_N	0.483151 (0.03849535)

AR: Test 2

n1 = 248	n2 = 46	n3 = 12	n4 = 45

P(coh)	0.27634
θ_3C	0.5685493 (0.03781467)
θ_2C	0.1693967 (0.04116373)
θ_1C	0.02464633 (0.02105362)
θ_N	0.2374076 (0.03017369)

AR: Test 3

n1 = 260	n2 = 39	n3 = 17	n4 = 35

P(coh)	0.90654
θ_3C	0.6226774 (0.03583543)
θ_2C	0.1207715 (0.04119916)
θ_1C	0.05714307 (0.03489953)
θ_N	0.199408 (0.03022842)

AR: Test 4

n1 = 293	n2 = 32	n3 = 8	n4 = 18

P(coh)	0.78936
θ_3C	0.735454 (0.03257557)
θ_2C	0.1302977 (0.03494289)
θ_1C	0.03227044 (0.0230069)
θ_N	0.1019778 (0.02161529)

Figure 5.5: Tables of Experiment 2's average parameter estimates for the EO model. The estimates are for the Adjacent-Related condition. The standard deviations for each estimate are shown in parentheses.

AU: Test 1

n1 = 123	n2 = 59	n3 = 42	n4 = 127

P(coh)	0.46214
θ_3C	0.1779112 (0.03604287)
θ_2C	0.07921886 (0.04444857)
θ_1C	0.04848764 (0.03723827)
θ_N	0.6943823 (0.04247991)

AU: Test 2

n1 = 121	n2 = 74	n3 = 41	n4 = 115

P(coh)	0.66198
θ_3C	0.1324788 (0.03875348)
θ_2C	0.1700067 (0.05601529)
θ_1C	0.06135627 (0.04403593)
θ_N	0.6361583 (0.04401854)

AU: Test 3

n1 = 214	n2 = 55	n3 = 21	n4 = 61

P(coh)	0.58382
θ_3C	0.4477034 (0.03995082)
θ_2C	0.1763043 (0.04652175)
θ_1C	0.04155239 (0.03159659)
θ_N	0.33444 (0.03529736)

AU: Test 4

n1 = 203	n2 = 46	n3 = 21	n4 = 81

P(coh)	0.16196
θ_3C	0.4422928 (0.03792709)
θ_2C	0.1049112 (0.04261875)
θ_1C	0.02779145 (0.02402124)
θ_N	0.4250046 (0.0372955)

Figure 5.6: Tables of Experiment 2's average parameter estimates for the EO model. The estimates are for the Adjacent-Unrelated condition. The standard deviations for each estimate are shown in parentheses.

NR: Test 1

n1 = 126	n2 = 72	n3 = 42	n4 = 111

P(coh)	0.7711
θ_3C	0.1518564 (0.038676)
θ_2C	0.1579764 (0.05589765)
θ_1C	0.0711136 (0.04801231)
θ_N	0.6190536 (0.04448796)

NR: Test 2

n1 = 141	n2 = 79	n3 = 49	n4 = 82

P(coh)	0.99532
θ_3C	0.1744167 (0.04093068)
θ_2C	0.1700999 (0.06244186)
θ_1C	0.1881438 (0.06294801)
θ_N	0.4673396 (0.04482424)

NR: Test 3

n1 = 229	n2 = 42	n3 = 33	n4 = 47

P(coh)	0.84798
θ_3C	0.52336 (0.03657234)
θ_2C	0.06420423 (0.03915551)
θ_1C	0.1422368 (0.04649788)
θ_N	0.270199 (0.03613651)

NR: Test 4

n1 = 225	n2 = 59	n3 = 21	n4 = 46

P(coh)	0.89202
θ_3C	0.4675195 (0.04050903)
θ_2C	0.2099094 (0.048575)
θ_1C	0.06229039 (0.03820425)
θ_N	0.2602808 (0.03387433)

Figure 5.7: Tables of Experiment 2's average parameter estimates for the EO model. The estimates are for the Near-Related condition. The standard deviations for each estimate are shown in parentheses.

NU: Test 1

n1 = 103	n2 = 60	n3 = 52	n4 = 136

P(coh)	0.58972
θ_3C	0.114073 (0.03390679)
θ_2C	0.05952549 (0.0403463)
θ_1C	0.0687919 (0.04528353)
θ_N	0.7576096 (0.04473339)

NU: Test 2

n1 = 121	n2 = 60	n3 = 45	n4 = 125

P(coh)	0.61886
θ_3C	0.1685665 (0.03640166)
θ_2C	0.07901806 (0.04539315)
θ_1C	0.06068506 (0.04270549)
θ_N	0.6917304 (0.04389326)

NU: Test 3

n1 = 156	n2 = 69	n3 = 36	n4 = 90

P(coh)	0.8351
θ_3C	0.2452443 (0.04047218)
θ_2C	0.177723 (0.05493919)
θ_1C	0.07287431 (0.04697721)
θ_N	0.5041584 (0.04245061)

NU: Test 4

n1 = 178	n2 = 61	n3 = 34	n4 = 78

P(coh)	0.9054
θ_3C	0.3298132 (0.04011168)
θ_2C	0.1476036 (0.05259036)
θ_1C	0.08242092 (0.04824477)
θ_N	0.4401623 (0.04151703)

Figure 5.8: Tables of Experiment 2's average parameter estimates for the EO model. The estimates are for the Near-Unrelated condition. The standard deviations for each estimate are shown in parentheses.

FR: Test 1

n1 = 102	n2 = 87	n3 = 49	n4 = 113

P(coh)	0.8092
θ_3C	0.05201878 (0.03120839)
θ_2C	0.2004677 (0.05888173)
θ_1C	0.1089305 (0.05944297)
θ_N	0.638583 (0.04710866)

FR: Test 2

n1 = 208	n2 = 56	n3 = 22	n4 = 65

P(coh)	0.55718
θ_3C	0.428249 (0.03970613)
θ_2C	0.1753613 (0.0467184)
θ_1C	0.04095528 (0.03115032)
θ_N	0.3554344 (0.03590397)

FR: Test 3

n1 = 198	n2 = 67	n3 = 30	n4 = 56

P(coh)	0.9815
θ_3C	0.3692225 (0.04171923)
θ_2C	0.2071416 (0.0543344)
θ_1C	0.1037961 (0.04905921)
θ_N	0.3198398 (0.03785477)

FR: Test 4

n1 = 261	n2 = 43	n3 = 17	n4 = 30

P(coh)	0.96122
θ_3C	0.6140322 (0.0369969)
θ_2C	0.1451741 (0.04294788)
θ_1C	0.06766266 (0.03644577)
θ_N	0.1731311 (0.02884536)

Figure 5.9: Tables of Experiment 2's average parameter estimates for the EO model. The estimates are for the Far-Related condition. The standard deviations for each estimate are shown in parentheses.

FU: Test 1

n1 = 136	n2 = 97	n3 = 28	n4 = 90

P(coh)	0.40116
θ_3C	0.1109607 (0.04182083)
θ_2C	0.3623125 (0.05632655)
θ_1C	0.04009803 (0.03270722)
θ_N	0.4866288 (0.03950195)

FU: Test 2

n1 = 184	n2 = 101	n3 = 26	n4 = 40

P(coh)	0.9961
θ_3C	0.2340874 (0.04604337)
θ_2C	0.4218371 (0.05997312)
θ_1C	0.1135169 (0.04650949)
θ_N	0.2305586 (0.03358945)

FU: Test 3

n1 = 205	n2 = 80	n3 = 17	n4 = 49

P(coh)	0.59234
θ_3C	0.352063 (0.04390091)
θ_2C	0.3418137 (0.05146091)
θ_1C	0.03750667 (0.02861785)
θ_N	0.2686167 (0.03248717)

FU: Test 4

n1 = 222	n2 = 83	n3 = 15	n4 = 31

P(coh)	0.89066
θ_3C	0.3915629 (0.0446187)
θ_2C	0.3796499 (0.05182039)
θ_1C	0.05206266 (0.03242427)
θ_N	0.1767246 (0.02858652)

Figure 5.10: Tables of Experiment 2's average parameter estimates for the EO model. The estimates are for the Far-Unrelated condition. The standard deviations for each estimate are shown in parentheses.

Experiment 3: Table of Pooled Frequencies

		Attempt			
		1st	2nd	3rd	4th
Probe Type	FO1	183	9	4	4
	FO2	181	8	8	3
	FO3	191	7	2	0
	FO4	196	4	0	0

Figure 5.11: Table of pooled frequencies for item content questions of Experiment 3.

NU: Test 1

n1 = 48	n2 = 42	n3 = 31	n4 = 79

P(coh)	0.49116
θ_3C	0.04633653 (0.0318822)
θ_2C	0.09293035 (0.05543625)
θ_1C	0.09536265 (0.0627819)
θ_N	0.7653705 (0.05928146)

NU: Test 2

n1 = 65	n2 = 34	n3 = 27	n4 = 74

P(coh)	0.50526
θ_3C	0.143308 (0.04584756)
θ_2C	0.07751643 (0.05231837)
θ_1C	0.07253603 (0.05252771)
θ_N	0.7066395 (0.05611857)

NU: Test 3

n1 = 61	n2 = 32	n3 = 30	n4 = 77

P(coh)	0.40566
θ_3C	0.1262693 (0.04385989)
θ_2C	0.05862628 (0.04416516)
θ_1C	0.07627235 (0.05319468)
θ_N	0.7388321 (0.05568053)

NU: Test 4

n1 = 87	n2 = 36	n3 = 29	n4 = 48

P(coh)	0.79998
θ_3C	0.2427399 (0.05060247)
θ_2C	0.09566023 (0.06118775)
θ_1C	0.1824523 (0.07370317)
θ_N	0.4791476 (0.05890466)

Figure 5.12: Tables of Experiment 3's average parameter estimates for the EO model. The estimates are for the Near-Unrelated condition. The standard deviations for each estimate are shown in parentheses.

FU: Test 1

n1 = 87	n2 = 47	n3 = 17	n4 = 49

P(coh)	0.59524
θ_3C	0.1963957 (0.05550838)
θ_2C	0.2707776 (0.07336875)
θ_1C	0.0656899 (0.05000069)
θ_N	0.4671369 (0.05211989)

FU: Test 2

n1 = 97	n2 = 38	n3 = 21	n4 = 44

P(coh)	0.90716
θ_3C	0.2882796 (0.05324906)
θ_2C	0.1639239 (0.070632)
θ_1C	0.1127347 (0.06539409)
θ_N	0.4350618 (0.05461677)

FU: Test 3

n1 = 100	n2 = 38	n3 = 21	n4 = 41

P(coh)	0.93554
θ_3C	0.3030174 (0.05345023)
θ_2C	0.1657703 (0.07106654)
θ_1C	0.1234963 (0.06683245)
θ_N	0.4077161 (0.05444512)

FU: Test 4

n1 = 107	n2 = 39	n3 = 15	n4 = 39

P(coh)	0.7119
θ_3C	0.3339846 (0.05466784)
θ_2C	0.2194337 (0.06815587)
θ_1C	0.06960872 (0.05005708)
θ_N	0.376973 (0.04940756)

Figure 5.13: Tables of Experiment 3's average parameter estimates for the EO model. The estimates are for the Far-Unrelated condition. The standard deviations for each estimate are shown in parentheses.

FO: Test 1

n1 = 183	n2 = 9	n3 = 4	n4 = 4

P(coh)	0.8627
θ_Si	0.85
θ_Fi	0.078
θ_Ni	0.072
θ_F2	0.591

FO: Test 2

n1 = 181	n2 = 8	n3 = 8	n4 = 3

P(coh)	0.4994
θ_Si	0.839
θ_Fi	0.091
θ_Ni	0.07
θ_F2	0.358

FO: Test 3

n1 = 191	n2 = 7	n3 = 2	n4 = 0

P(coh)	0.92867
θ_Si	0.901
θ_Fi	0.084
θ_Ni	0.015
θ_F2	0.598

FO: Test 4

n1 = 196	n2 = 4	n3 = 0	n4 = 0

P(coh)	0.7228
θ_Si	0.94
θ_Fi	0.05
θ_Ni	0.09
θ_F2	0.716

Figure 5.14: Tables of Experiment 3's average parameter estimates for the IS model. The estimates are for the item content/foil condition.

Bibliography

Block, R. A. (1986). Remembered duration: Imagery processes and contextual encoding. *Acta Psychologica*, 62(2):103–122. [2]

Bower, G. H. (1961). Application of a model to paired-associate learning. *Psychometrika*, 26(3):255–280. [44]

Brown, G. D. A., Neath, I., and Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, 114(3):539–576. [3, 39]

Brown, G. D. A., Preece, T., and Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107(1):127–181. [3, 39]

Bush, R. R. and Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, 58(5):313–323. [45]

Cam, L. L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer New York. [13]

Chechile, R. A. (1998). A new method for estimating model parameters for multinomial data. *Journal of Mathematical Psychology*, 42(4):432–471. [9]

Chechile, R. A. (2004). New multinomial models for the chechile–meyer task. *Journal of Mathematical Psychology*, 48(6):364–384. [9, 12]

Chechile, R. A. (2009). Pooling data versus averaging model fits for some prototypical multinomial processing tree models. *Journal of Mathematical Psychology*, 53(6):562–576. [14, 15, 16]

Chechile, R. A. (2010a). Modeling storage and retrieval processes with clinical populations with applications examining alcohol-induced amnesia and korsakoff amnesia. *Journal of Mathematical Psychology*, 54(1):150–166. [14, 16]

Chechile, R. A. (2010b). A novel bayesian parameter mapping method for estimating the parameters of an underlying scientific model. *Communications in Statistics - Theory and Methods*, 39(7):1190–1201. [9, 10, 12]

Chechile, R. A. (2018). *Analyzing Memory*. MIT Press. [3, 4, 5, 39]

Chechile, R. A. (2020). *Bayesian Statistics for Experimental Scientists*. MIT Press. [46]

Chechile, R. A. and Pintea, G. I. (2021). Measuring components of the memory of order. *Journal of Mathematical Psychology*, 100:102476. [14, 17]

Chechile, R. A. and Sloboda, L. N. (2014). Reformulating markovian processes for learning and memory from a hazard function framework. *Journal of Mathematical Psychology*, 59:65–81. [40]

Chechile, R. A., Sloboda, L. N., and Chamberland, J. R. (2012). Obtaining separate measures for implicit and explicit memory. *Journal of Mathematical Psychology*, 56(1):35–53. [5, 12]

Chechile, R. A. and Soraci, S. A. (1999). Evidence for a multiple-process account of the generation effect. *Memory*, 7(4):483–508. [5]

DeCarlo, L. T. (2012). On a signal detection approach to -alternative forced choice with bias, with maximum likelihood and bayesian approaches to estimation. *Journal of Mathematical Psychology*, 56(3):196–207. [12]

Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of Neurosciences*, 20(4). [38]

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews Neuroscience*, 15(11):732–744. [3]

Estes, W. K. (1972). Elements and patterns in diagnostic discrimination learning. *Transactions of the New York Academy of Sciences*, 34(1 Series II):84–95. [2]

Estes, W. K. (1985). Memory for temporal information. In *Time, Mind, and Behavior*, pages 151–168. Springer Berlin Heidelberg. [2]

Fozard, J. L. (1970). Apparent recency of unrelated pictures and nouns presented in the same sequence. *Journal of Experimental Psychology*, 86(2):137–143. [2]

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, 113(1):44–66. [4, 37]

Hicks, R. E., Miller, G. W., and Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *The American Journal of Psychology*, 89(4):719. [2]

Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychological Review*, 77(3):223–233. [ii, 2, 37]

Hinrichs, J. V. and Buschke, H. (1968). Judgment of recency under steady-state conditions. *Journal of Experimental Psychology*, 78(4, Pt.1):574–579. [2]

Howard, M. W. and Eichenbaum, H. (2015). Time and space in the hippocampus. *Brain Research*, 1621:345–354. [3, 39]

Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., and Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of Neuroscience*, 34(13):4692–4707. [3]

Howard, M. W., Shankar, K. H., Aue, W. R., and Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, 122(1):24–53. [3, 39]

Jeffreys, S. H. (1998). *Theory of Probability*. OUP Oxford. [21]

Jenkins, L. J. and Ranganath, C. (2010). Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *Journal of Neuroscience*, 30(46):15558–15565. [3]

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795. [21]

Kesner, R. P., Hopkins, R. O., and Fineman, B. (1994). Item and order dissociation in humans with prefrontal cortex damage. *Neuropsychologia*, 32(8):881–891. [3]

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86. [15]

Lee, C. L. (1992). *Essays in honor of William K. Estes, Vol. 1. From learning theory to connectionist theory.*, chapter The perturbation model of short-term memory: A review and some further developments., pages 119–141. Lawrence Erlbaum Associates, Inc. [2, 61]

Lee, C. L. and Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning & Memory*, 7(3):149–169. [2, 61]

MacDonald, C. J., Lepage, K. Q., Eden, U. T., and Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, 71(4):737–749. [3]

Mangels, J. A. (1997). Strategic processing and memory for temporal order in patients with frontal lobe lesions. *Neuropsychology*, 11(2):207–221. [3]

Michon, J. A. and Jackson, J. L. (1984). Attentional effort and cognitive strategies in the processing of temporal information. *Annals of the New York Academy of Sciences*, 423(1 Timing and Ti):298–321. [2]

Milner, B., Corsi, P., and Leonard, G. (1991). Frontal-lobe contribution to recency judgements. *Neuropsychologia*, 29(6):601–618. [3]

Naya, Y. and Suzuki, W. A. (2011). Integrating what and when across the primate medial temporal lobe. *Science*, 333(6043):773–776. [3]

O'Keefe, J. and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3(3):317–330. [3]

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, 321(5894):1322–1327. [3]

Seelig, S. A., Rabe, M. M., Malem-Shinitski, N., Risse, S., Reich, S., and Engbert, R. (2020). Bayesian parameter estimation for the SWIFT model of eye-movement control during reading. *Journal of Mathematical Psychology*, 95:102313. [12]

Shimamura, A. P., Janowsky, J. S., and Squire, L. R. (1990). Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia*, 28(8):803–813. [2, 3]

Underwood, B. J. (1977). *Temporal Codes for Memories*. Lawrence Erlbaum Associates. [ii, 2, 34]

Underwood, B. J. (1983). Attributes of memory. *Psychological Review*, 76(6):559–573. [2]

Ward, R. D., Gallistel, C., and Balsam, P. D. (2013). It's the information! *Behavioural Processes*, 95:3–7. [3, 39]

Wearden, J. and Towse, J. (1994). Temporal generalizations in humans: Three further studies. *Behavioural Processes*, 32(3):247–263. [3]

Wearden, J. H. and Bray, S. (2001). Scalar timing without reference memory? episodic temporal generalization and bisection in humans. *The Quarterly Journal of Experimental Psychology B*, 54(4):289–309. [3]