INSIGHTS FROM MEMORY MODELS ON FALSE MEMORY PROCESSES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Koyuki Nakamura

December 2018

INSIGHTS FROM MEMORY MODELS ON FALSE MEMORY PROCESSES

Koyuki Nakamura, Ph. D.

Cornell University 2018

One of the main goals of memory research is to identify why false memory errors occur. Much work has been done in prior research to identify and define false memory processes using indirect measurements for them, such as using response latencies to hypothesize how one memory process may be faster over another, or inferring how recollecting contextual details of a target is affected by the order in which items were presented during encoding. A weakness in these indirect methods of measuring processes is just that: they are indirect measurements. The advantage of memory models is that they provide direct measurements of memory processes, allowing researchers to test assumptions about how manipulations affect them, and comparing them against each other.

The theme of this dissertation is to use new models and modeling techniques to answer questions that previously relied on indirect methods of identifying processes. In Experiment 1, a new model was created to accommodate two source memory designs used to measure overdistribution in episodic memory. The model was able to provide insight that the two designs tapped into different memory processes, and thus were not measuring the same overdistribution metric as previously assumed. In Experiment 2, a new methodology in modeling was implemented so that relative process speeds could be measured alongside process parameter estimates. The new latency extension of a widely used recognition model was fit to previously collected data, providing a first look into how familiarity may not be the fast process responsible for false memory under fast response deadlines as previously believed. Experiment 3 addressed

weaknesses from Experiment 2, and supported its findings that context recollection was faster than familiarity. These experiments demonstrate how memory models provide a simple but powerful tool to answer questions that could only be inferred previously.

BIOGRAPHICAL SKETCH

Koyuki was a double major in Psychology and Biology with a concentration in neuroscience at Cornell University, and graduated with a Magna Cum Laude for her bachelor's degree in 2012. During her undergraduate education, she worked in Dr. Charles Brainerd's lab of Memory and Neuroscience, involving herself in projects that would lay the foundation for her modeling work during her graduate career. Her work centered on memory overdistribution errors: a newly identified type of memory distortion different from standard false memory errors.

As a PhD student, she continued her work on overdistribution errors. She worked on a series of projects that helped develop a better understanding on the mechanisms of conjunction and disjunction fallacies in episodic memory, and their relation to overdistribution. She was also involved in a wide breadth of other projects, including a study looking at different memory processes involved in categorical memory judgments versus confidence ratings, studies looking at the relationship between biomarkers that predict Alzheimer's dementia and memory processes, studies testing recollection as a bivariate rather than a univariate process, and a study using math modeling techniques to identify brain regions responsible for true and false memory processes. Her dissertation is a product of her training in memory modeling.

ACKNOWLEDGMENTS

I would like to thank Dr. Brainerd, Dr. Reyna, and Dr. Ceci for their support during my doctorate program. Their mentorship was invaluable, and I learned a lot from them not only about research, but also life lessons I will carry for many years. I grew during the program thanks to their outstanding mentorship.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

It has not always been the case that false memory was something readily accepted as a common occurrence. Indeed, the public attitude that memory is infallible, especially when the person remembering them has high confidence in its accuracy, is reflected in how conviction outcomes of criminal cases are often dependent on eyewitness testimony (Fisher, 1997). However, data show that out of 356 wrongful convictions exonerated through DNA evidence, 70% of the convictions involved an eyewitness misidentification, suggesting that witnesses not only did not accurately remember the perpetrator, but had confidence in their false memory (Innocence Project, 2016). Why do false memory errors occur, and in what situations are they likely to appear? The more we understand memory errors, the more we can identify the likelihood that a memory decision is accurate.

To answer these two questions, it is first necessary to define what false memory is. Traditionally, three types of memory responses have been defined: a.) True memory, where one remembers an event exactly as it happened, b.) Forgetting, where one cannot recall anything about an event, and c.) False memory, where one remembers details of an event that never happened while rejecting true details. Recently, it has been proposed that there is another type of memory error called overdistribution. Whereas false memory is accepting a false detail and rejecting a true detail (e.g. remembering the culprit was Caucasian and not Hispanic), overdistribution is when a person accepts both the false and true details even in cases where the details contradict each other (e.g. remembering the culprit as being both Caucasian and Hispanic) (Brainerd, Reyna, Holliday, & Nakamura, 2012). The way memory is encoded and retrieved determine the probability of true memory, forgetting, false memory, and overdistribution, and

they are defined by memory theories.

One prominent theory that defines these different types of memory responses is fuzzy trace theory (FTT; Reyna & Brainerd, 2011). FTT defines two independent memory traces that contribute to true and erroneous memory retrieval; verbatim and gist. Verbatim traces are the surface level features of an experience; for example remembering you had fries for lunch because you could remember the texture and color of the fries, and their oily smell. Because verbatim traces are traces that exemplify the event as it had actually happened, they support true memory. Their weakness, however, is that they tend to be prone to interference, and decay relatively quickly over time. Gist traces, on the other hand, are traces that preserve the underlying meaning of an event. In the prior example, you may remember that you had fries for lunch because you could remember that you had fast food. Because gist traces can be consistent with many different episodic states of an event (e.g. "fast food" can define both fries and burgers), it not only supports true memory (remembering 'fries') but also erroneous memory (remembering 'burgers'). Therefore, while these traces are more robust than verbatim traces, they can be relatively error-prone.

The strength of FTT is that its simple framework of two independent memory traces can not only explain memory phenomena, but also predict them. For example, it predicts why false memory is especially elevated in a widely used false memory paradigm, the Deese-Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1985). In this paradigm, participants study a list of words where some words are related to each other. They then take a recognition test where they are presented with words that they had previously studied (targets), new words that were related to the studied targets (critical distractors), and new words that were unrelated to the studied targets (unrelated distractors). One of the reasons why this paradigm is

widely used is because it produces reliable and high levels of false memory to the critical distractors. Looking at the FTT framework, it is easy to see why this would be so. Because critical distractors share the same gist traces as the studied targets, a memory decision based on gist would support not only the acceptance of targets, but also critical distractors. Interestingly, critical distractors may be accepted with high levels of confidence, and is sometimes accompanied by vivid imagery of its presentation even though it was never presented (Brainerd, Wright, Reyna, & Mojardin, 2001).

In another example, FTT was the first theory to predict the overdistribution error. As described before, overdistribution is a memory error where not only is the true event accepted, but the false event is simultaneously accepted even in cases where the true and false events are mutually exclusive. Overdistribution usually requires an indirect method of measurement, as asking someone if an item is both a target and a critical distractor, for example, may be rejected simply on a metacognitive basis rather than memory (e.g. something can't be both a target and a critical distractor based on definitions alone). Designs that measure overdistribution rely on the conjoint recognition paradigm (Brainerd et al., 2001). In this paradigm, participants are given a standard recognition task but with some new probes during test. In a standard design, participants are usually asked if an item was 'old' or 'new'. In conjoint recognition, participants are asked three questions: "Was it a target", "Was it new but related to a target", and "Was it either a target or new but related to a target." These three probes can then be used to distinguish if a memory error was due to forgetting, false memory, or overdistribution. Let us take the example where Bagpipe was a critical distractor in a DRM task. In the case of forgetting, all three probes would be rejected. In the case of false memory, the probe "it was a target" and "it was either a target or new but related to a target" would be accepted and "it was new but related"

would be rejected. In the case of overdistribution, all three probes would be accepted. Again, it is easy to see how overdistribution can be predicted in semantic false recognition paradigms with FTT's definition of gist traces: because gist traces support all three probes in the conjoint recognition paradigm, it supports acceptances for all three descriptions, resulting in overdistribution.

The purpose of this dissertation is to improve upon the existing memory models to better represent memory processes. There were two sets of experiments that targeted two different gaps in the current memory modeling literature. The first builds upon a series of studies that have looked at overdistribution errors. Because overdistribution errors are a relatively new discovery, there currently does not exist a model that sufficiently accommodates the two methods, direct and indirect, in which it is measured. The second set of experiments address a weakness in existing models, in which they are not able to measure the relative speed of memory processes. Because the first and second set of experiments focus on different models and different experimental designs, I will first discuss the modeling background of overdistribution errors followed by details of Experiment 1, and then I will discuss the modeling background involved in measuring process speeds followed by details of Experiment 2 and Experiment 3.

Overdistribution Errors

Because overdistribution errors are defined by their contradictory nature of an item as having been remembered in both correct and incorrect contexts, it is often impractical to measure its occurrence directly (e.g. asking a subject if a test item is both a target and a related distractor). In the earliest research on overdistribution errors, the phenomenon was measured in semantic false memory designs using indirect measures. With the conjoint recognition design, these studies had available three independent estimates of responses for each test cue, which could be

used to calculate overdistribution: (a) P(T), the probability that an item is judged to be a target; (b) P(R), the probability that an item is judged to be a related distractor; and (c) P(TUR), the probability that an item is judged to be either a target or a related distractor.

Taking advantage of Kolmogorov's axiom, a simple rule of probability, it is possible to estimate the overdistribution index $P(\mathrm{T} \cap \mathrm{R})$, or the probability that an item is judged to be both a target and a related distractor. For any two events T and R, the sum of their probabilities equal the sum of their disjunction and their conjunction:

$$P(\mathrm{T}) + P(\mathrm{R}) = P(\mathrm{TUR}) + P(\mathrm{T} \cap \mathrm{R}) \qquad (1)$$

Written another way, it is:

$$P(\mathrm{T}) + P(\mathrm{R}) - P(\mathrm{TUR}) = P(\mathrm{T} \cap \mathrm{R}) \qquad (2)$$

Logically, the probability of the conjunction must equal zero, because it is impossible for something to be both a target and a related distractor. A positive value of $P(\mathrm{T} \cap \mathrm{R})$, then, is the measure of overdistribution.

Brainerd & Reyna (2008) first used FTT to predict overdistribution, and Brainerd et al. (2010) first formalized it in a mathematical model and demonstrated how manipulations that increased gist retrieval increased the memory distortion. Brainerd & Reyna (2008) calculated overdistribution by compiling data of the three conjoint recognition response probabilities in a between subjects design. Though they observed evidence of overdistribution as FTT predicted, the weakness was that, because it was a between-subjects design where no one subject received all three conjoint recognition probes, it was impossible to test for it in individual subjects. Brainerd et al. (2010) addressed this in a design where participants received all three conjoint recognition probes in a single experiment, and they were able to show how manipulations that increased reliance on gist traces (such as time delay between study and test) increased

overdistribution, again as predicted by FTT.

In 2012, Brainerd et al. showed how overdistribution errors are not only observed in simple recognition experiments but are also prevalent in source-monitoring designs. In the source-monitoring designs, subjects are presented with two (or more) lists of words with different font and background color, and are asked to make source judgments rather than recognition judgments. In a two-list design, then, the conjoint recognition questions are: (a.) "it was presented on List 1"; (b) "it was presented on List 2"; (c) "it was presented in List 1 or List 2." In these experiments, none of the words were ever presented on both lists, so similar to the semantic design, the probability of an item as having appeared on both lists, $P(L_1 \cap L_2)$ was zero. Not only did these experiments show that overdistribution is ubiquitous in item recognition but also source monitoring, they provided a way to directly measure the overdistribution index $P(L_1 \cap L_2)$. By definition, it is not impossible for something to have been presented on both lists, unlike defining something as being both a target and a related distractor. However, experimentally, it can be constrained so that the conjunctive probe is impossible. Therefore, we can have a situation in which participants are asked about the conjunctive probe without having to worry about non-memorial, metacognitive judgments that would lead to an automatic rejection of the probe.

FTT predicts overdistribution in semantic false memory designs, but also in source designs as well. Verbatim traces and gist traces support two kinds of recollections that are important when making source judgments: target and context recollection (Brainerd, Gomes, & Moran, 2014). Target recollection is when traces of an item's presentation can be retrieved, while context recollection is when traces of the contextual details surrounding an item's presentation can be retrieved. Until recently, it has been assumed that context recollection is

intrinsically linked to target recollection, in that context recollection occurs only if there is target recollection. However, there is growing evidence supporting the idea that not only can one have target recollection without context recollection, but that they can have context recollection without target recollection (e.g. Ceci, Fitneva, & Williams, 2010; Brainerd et al., 2014; Brainerd, Gomes, & Nakamura, 2015). Overdistribution in source memory, then, increases when there is target recollection without context recollection, but is suppressed when there is context recollection regardless of whether or not there is target recollection. This is what was found in source memory designs (Brainerd et al., 2012; Nakamura & Brainerd, 2016).

Here, it is worth noting that these indirect and direct methods of measuring overdistribution are similar to phenomena found in the judgment and decision-making literature. The indirect overdistribution index is analogous to the disjunction fallacy. The disjunction fallacy occurs when the sum of the probability of two mutually exclusive events (e.g. dying of a heart attack and dying of lung disease) is judged to be greater than the probability of their disjunction (e.g. dying of *either* a heart attack *or* lung disease), and this fallacy was first observed by Tversky and Koehler (1994).

In the case of the direct measure of overdistribution, two types are discussed in the memory literature. Conjunction illusions, which is when the impossible conjunction (e.g. judging something is both a target and a related distractor) is accepted at a probability greater than zero. In some cases, these conjunction illusions can be so erroneous that they elevate to the level of conjunction *fallacies*, where the impossible conjunction is judged to be more probable than either of its constituents. The conjunction fallacy is a famous example in judgment and decision-making in the form of Tversky and Kahneman's (1983) Linda problem. In the classic Linda problem, participants read a prompt about a left-leaning woman named Linda, and then are

asked to make three probability judgments: a.) the probability that Linda is a bank teller; b.) the probability that Linda is a feminist; c.) the probability that Linda is a bank teller *and* a feminist. Based on rules of probability, it is impossible for it to be more probable that Linda is both a bank teller and a feminist than she is to be just a bank teller, but this logical fallacy commonly occurs. A memory conjunction fallacy in a source memory design occurs when $P(L_1 \cap L_2) > P(L_1)$ or $P(L_1 \cap L_2) > P(L_2)$, and it has been observed in prior studies (Brainerd, Nakamura, Reyna, & Holliday, 2017; Nakamura & Brainerd, 2016).

While this is not in the scope of the focus of the dissertation, the fact that memory analogues to decision fallacies exist suggests that the causes of these fallacies may originate from basic encoding and retrieval processes; that memory and judgment are linked. This idea is not very surprising considering how fallacies in judgment and decision-making have been explained by assumptions in memory. For example, support theory (ST), which was first used to describe disjunction fallacies by Tversky & Koehler (1994), assumes that the fallacies arise from limitations in cognitive resources. According to ST, disjunctive probes require more mental resources to unpack into their individual components, and therefore require more effort to retrieve evidence representative of them which in turn leads to an underestimation of the disjunctive probe. A link between judgment and decision-making and memory seems obvious, and yet very few studies have been conducted in exploring the link between fallacies in judgment and decision-making and fallacies in memory. Because the designs to study overdistribution are memory analogues to two fallacies in judgment and decision-making, a study in overdistribution may provide some unique insights into process-level mechanisms of the fallacies in decision making as well.

What is unique about FTT is that it predicts overdistribution, and stipulates conditions in

which it will increase (e.g. manipulations that increase target recollection), rather than simply explain why it might exist. The source guessing model by Batchelder and Riefer (1990), is one model that can accommodate overdistribution, but does not predict it the way FTT is able to. The source guessing model is defined over a source recognition task where subjects study two (or more) lists of words. During test they are first asked to make an old-new recognition judgment, and if an item is judged to be old, they make a source judgment about it. In this model, there are three memory parameters (D_1, d_1, D_N), one guessing parameter (g), and one bias parameter (b). D_1 is the probability that subject has target recollection for a List 1 target, d_1 is the probability that, for words where participants had target recollection for a List 1 target, they have context recollection, and D_N is the probability that subjects correctly judge distractors as being new. The guessing parameter g is the probability that for List 1 items where target recollection could be retrieved, subjects guess its source as the one mentioned in the test probe, and b is the probability that for distractors, subjects guess its source as the one mentioned in the test probe. For a List 1 item, subjects will accept the correct probe $P(L_1)$ when they have target recollection and context recollection, and also when they only have target recollection and they guess its context. They will accept the incorrect probe $P(L_2)$ when they only have target recollection and they guess its context, and they will accept the disjunctive probe $P(L_1 U L_2)$ when they have target and context recollection, or when they only have target recollection and guess its context. As you can see with these definitions, whether or not there is overdistribution is dependent on the value of the guessing parameter, because when g=0.5 the relationship of $P(L_1)$, $P(L_2)$, and $P(L_1 U L_2)$ becomes additive. However, because overdistribution in the source guessing model is entirely dependent on a non-memorial process 'guessing,' it is not possible to make predictions about conditions that will increase or decrease the phenomenon, and by

extension it will not be able to explain how it can be brought under experimental control.

Disjunction and Conjunction Fallacies in Memory

The main focus of Experiment 1 is on the source design of overdistribution, rather than the semantic false memory design. Source designs provide a unique opportunity to have all four response probability estimates measured directly from the experimental design, so we have all four estimates from the equation $P(\mathrm{T}) + P(\mathrm{R}) - P(\mathrm{TUR}) = P(\mathrm{T} \cap \mathrm{R})$. In other words, it is possible to see how well the indirect measurement of overdistribution $P(\mathrm{L}_1) + P(\mathrm{L}_2) - P(\mathrm{L}_1 \mathrm{U} \mathrm{L}_2)$ predicts the direct measure of overdistribution $P(\mathrm{L}_1 \cap \mathrm{L}_2)$.

What would this allow us to accomplish? First, this equality assumes that the two designs used to study overdistribution are inherently linked, and we can test that assumption. Second, this equality also assumes that the two fallacies embedded in these designs, disjunction and conjunction fallacies, are also linked to each other. Prior experiments that have studied memory disjunction and conjunction fallacies separately, however, have shed light that methods that measure overdistribution directly or indirectly do not tap the same memory processes. In Brainerd et al. (2014), direct measures of overdistribution were compared with indirect measures of overdistribution in the same general design, but in two different experiments. In the design, participants studied two or three lists of uncategorized words that differed in frequency and concreteness, and took either a disjunction fallacy source memory experiment (indirect overdistribution measure) or a conjunction fallacy source memory experiment (direct overdistribution measure). In general, the estimates of indirect and direct overdistribution measures differed reliably.

A separate study in 2016 by Nakamura and Brainerd first showed how memory processes may differ for indirect and direct measures of overdistribution, or disjunction and conjunction

fallacies in memory. In their design, subjects studied two lists of words, some words belonging in categories while other words were not. Then, they took a source recognition test with either disjunctive or conjunctive probes. Because this design used categorized words and therefore test probes not only included targets but also related distractors, it was possible to test whether or not disjunction and conjunction fallacies could occur for items that had never been presented. Related distractors were a test to see if target recollection was an important factor in the occurrence of these fallacies. The data suggested that target recollection was important, but only for disjunction fallacies. Disjunction fallacies were only observed for targets, lending evidence to the fact that target recollection is important for these fallacies to occur. Similarly, they only appeared for List 1 targets than for List 2 targets, and it has been shown previously that target recollection is susceptible to proactive interference and therefore would be greater for List 1 items (Brainerd, Gomes, & Nakamura, 2015). Unlike disjunction fallacies, conjunction illusions appeared for both targets and related distractors, lending evidence to the fact that target recollection is not necessary for these errors to occur. Furthermore, conjunction fallacies were greater for List 2, rather than List 1, suggesting that not only are they not reliant on target recollection, but may be reliant on context recollection.

It is clear that there seems to be process differences for memory disjunction and conjunction fallacies. However, it is not currently possible to directly compare such processes because there is no model that accommodates both disjunction and conjunction designs. While a few models have been proposed for disjunction fallacy designs, but there is no memory model that fits conjunction fallacy designs to date, and one that also accommodates disjunction fallacies. I will briefly discuss two models that have been used to fit memory disjunction fallacy designs, and then discuss how one of the models has been modified for Experiment 1 to include

conjunction fallacies.

CPD and DR Model

The conjoint process dissociation (CPD) model was first developed by Brainerd et al. (2012) to fit disjunction fallacy data in source memory designs. The model was a modification of Jacoby's (1991) process-dissociation model, which uses the same source recognition design often used in source overdistribution studies, but without the disjunctive probe. The process-dissociation model was defined over acceptance probabilities for targets presented on List 1 as follows:

$$P(\mathrm{L}_1?|\mathrm{L1}) = R + (1 - R)F \qquad (3)$$

$$P(\mathrm{L}_2?|\mathrm{L1}) = (1 - R)F \qquad (4)$$

In the model, R is the probability that a List 1 target's presentation can be recollected, and F is the probability that a List 1 target's presentation cannot be recollected but is accepted on the basis of familiarity.

The CPD model built upon this so that there were enough degrees of freedom to separate overdistribution from true and false source memory with the additional conjoint recognition probe $\mathrm{L}_1\mathrm{UL}_2$?. For List 1 targets, the model is as follows:

$$P(\mathrm{L}_1?|\mathrm{L}_1) = R_1 + (1 - R_1)E_2 + (1 - R_1)(1 - E_2)F_1 \qquad (5)$$

$$P(\mathrm{L}_2?|\mathrm{L}_1) = (1 - R_1)E_2 + (1 - R_1)(1 - E_2)F_1 \qquad (6)$$

$$P(\mathrm{L}_1\mathrm{UL}_2?|\mathrm{L}_1) = R_1 + (1 - R_1)(1 - E_2)F_1 \qquad (7)$$

R_1 is the probability that a List 1 cue prompts the retrieval of List 1 contextual details, E_2 is the probability that a List 1 cue prompts the erroneous retrieval of List 2 contextual details, and F_1 is the probability that a List 1 cue prompts item memory.

While the CPD model has fit data from source disjunction fallacy studies well, it has a

weakness in how it is defined. In particular, the *R* parameter is not able to distinguish acceptances based on target recollection and context recollection. In the CPD model, "recollection" might be referring to the retrieval of details surrounding the target's presentation, or the retrieval of details surrounding the context of the target's presentation. The dual recollection (DR) model, on the other hand, gets around this problem by redefining the parameters to incorporate target recollection, context recollection, and familiarity. Of course there is a concern that familiarity is simply something derived from context recollection, and that rather than having a bivariate recollection model in addition to familiarity, data can be accounted for with either just the target-context recollection distinction, or the recollection-familiarity distinction. Brainerd, Gomes, & Moran (2015) introduced different experimental manipulations that should affect target recollection, context recollection, and familiarity differently, in order to see if they were distinct from each other. They used a source memory conjoint recognition design, with key manipulations being List order (predicted to affect target and context recollection in opposite directions), word frequency (predicted to affect target recollection but not context recollection), and word concreteness (predicted to affect frequency, but not either forms of recollection). The model fit their data in addition to providing evidence that the three key manipulations affected the two kinds of recollection and familiarity in different ways. Therefore, the DR model is much more comprehensive than the CPD model.

For List 1 targets, the DR model is as follows:

$$P(\mathrm{L}_1?|\mathrm{L}_1) = C_1 + (1 - C_1)T_1 + (1 - C_1)(1 - T_1)b \quad (8)$$

$$P(\mathrm{L}_2?|\mathrm{L}_1) = (1 - C_1)T_1 + (1 - C_1)(1 - T_1)b \quad (9)$$

$$P(\mathrm{L}_1\mathrm{UL}_2?|\mathrm{L}_1) = C_1 + (1 - C_1)T_1 + (1 - C_1)(1 - T_1)F_1 + (1 - C_1)(1 - T_1)(1 - F_1)b_{12} \quad (10)$$

C_1 is the probability that a List 1 cue prompts retrieval of contextual details surrounding its

presentation. T_1 is the probability that a List 1 cue prompts a conscious reinstatement of its presentation on the study list. F_1 is the probability that a List 1 cue invokes a high level of familiarity that leads to its acceptance as having been presented at study. The two bias parameters b and b_{12} are the probability of accepting the cue based on non-memorial tendencies to accept a cue.

Revised DR Model

The DR model was chosen as a candidate to modify in a way to accommodate the conjunction fallacy source memory experiment, so that memory processes can be measured and compared between memory disjunction and conjunction fallacies. Everything is the same as the standard DR model, except for the addition of the conjunction probe, and a new memory parameter. For List 1 targets, the modified DR model is as follows:

$$P(\mathrm{L}_1?|\mathrm{L}_1) = C_1 + (1 - C_1)T_1 + (1 - C_1)(1 - T_1)b \tag{11}$$

$$P(\mathrm{L}_2?|\mathrm{L}_1) = (1 - C_1)T_1 + (1 - C_1)(1 - T_1)P_1 + (1 - C_1)(1 - T_1)(1 - P_1)b \tag{12}$$

$$P(\mathrm{L}_1\mathrm{U}\mathrm{L}_2?|\mathrm{L}_1) = C_1 + (1 - C_1)T_1 + (1 - C_1)(1 - T_1)F_1 + (1 - C_1)(1 - T_1)(1 - F_1)b_{12} \tag{13}$$

$$P(\mathrm{L}_1 \cap \mathrm{L}_2?|\mathrm{L}_1) = C_1P_1 + C_1(1 - P_1)b_{1\cap 2} + (1 - C_1)\, b_{1\cap 2} \tag{14}$$

The new parameter, P, is phantom context recollection, and is the probability that a cue invokes the recollection of erroneous contextual details consistent with the wrong source, and the full list of the revised dual recollection model parameters can be found on Table 1.1. It is analogous to the phantom recollection parameter in the conjoint recognition model for standard recognition designs (Brainerd et al., 2001). Phantom recollection is a phenomenon observed in semantic false memory paradigms, where subjects retrieve erroneous contextual details for items that were never studied. This phenomenon is known to increase with manipulations that strongly increase

retrieval of gist traces.

Table 1.1

Parameters of the Revised Dual Recollection Model

Parameter	Definition
C_1	Probability that a List 1 cue prompts retrieval of contextual details surrounding its presentation
T_1	Probability that a List 1 cue prompts a conscious reinstatement of its presentation on the study list
F_1	Probability that a List 1 cue invokes a high level of familiarity that leads to its acceptance as having been presented at study
P_1	Probability that a List 1 cue invokes the recollection of erroneous contextual details consistent with List 2
b	Probability of accepting a non-disjunctive/conjunctive cue based on bias
b_{12}	Probability of accepting a disjunctive cue based on bias
$b_{1 \cap 2}$	Probability of accepting a conjunctive cue based on bias

Phantom recollection was first proposed when it was found that while theories without it could explain false alarms on the basis of familiarity, it could not explain false alarms that were accompanied by details vivid enough as though subjects could re-experience their presentation. Brainerd et al. (2001) demonstrated improved model fits when phantom recollection was included in the conjoint recognition model, and their experiments showed that it increased with retrieval of strong gist traces, suggesting that strong gist encouraged retrieval of contextual cues surrounding a related target's presentation. On this basis, it is possible for something similar to

happen in source monitoring studies as well. For example, imagine a study in which a subject is asked to study two lists of words with shared categories across both lists. A word (e.g. oak) presented on List 1 was never presented on List 2, but the gist of it belonging in a category (e.g. trees) was shared between both lists. Similar to phantom recollection, if strong gist can encourage retrieval of contextual cues surrounding a related target's (e.g. maple) presentation, then it seems plausible that in some cases it may lead to the retrieval of erroneous contextual cues from the wrong source (e.g. retrieving List 2 contextual cues for a List 1 word "oak" because related word "maple" was on List 2).

This is the basis for phantom context recollection in the modified dual recollection model. This is also similar to the *E* parameter in the CPD model, and thus this revised model can be thought of as a combination of both the original DR model and the CPD model. The addition of this new parameter changes two parts to the original model: accepting the wrong probe for a cue (e.g. accepting L_2? for a List 1 target), and accepting the new, conjunctive probe added to it ($L_1 \cap L_2$?). There is no change for $P(L_1?|L_1)$, because phantom recollection (and phantom context recollection) occurs when contextual cues are pulled from the probe, so that contextual cues from the wrong source will not be pulled unless it is mentioned in the probe (e.g. $L_2?|L_1$). In addition, the parameter is not included in the disjunctive probe, because the disjunctive probe is accepted if traces consistent with either List 1 or List 2 are retrieved, and the effect of accepting on the basis of phantom context recollection should be negligible compared to accepting on the basis of the other memory processes. However, it is an important piece for accepting conjunctive probes. Conjunctive probes require subjects to remember traces that are consistent with not only the correct source, but also consistent with the incorrect source. Therefore, conjunctive probes are accepted when there is both context recollection and phantom context recollection, but is

dependent on bias when there is only one or the other. The tree for the revised dual recollection model is available on Figure 1.1.

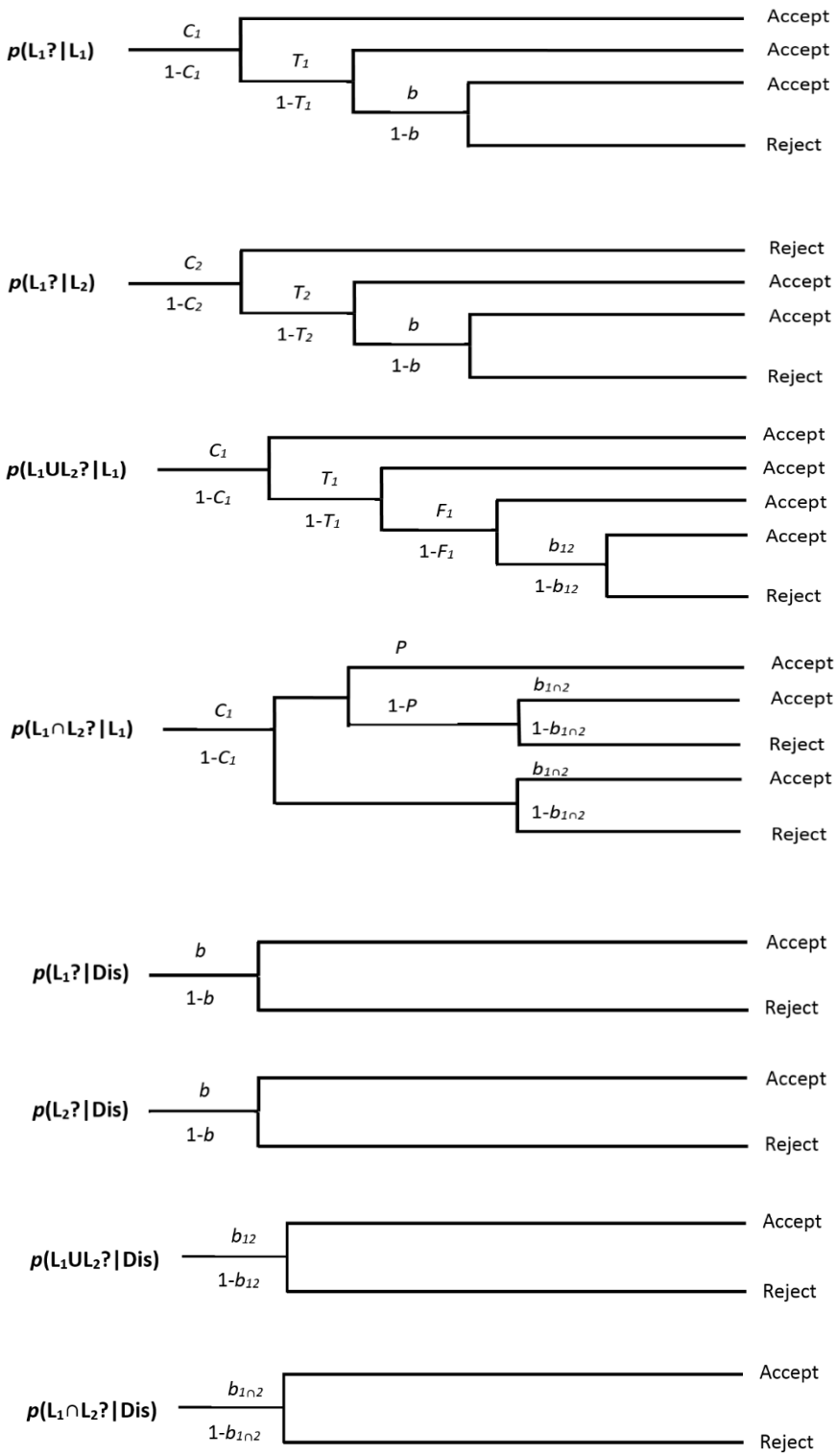


Figure 1.1. Revised dual recollection model tree for targets and distractors.

CHAPTER 2

EXPERIMENT 1: MODELING CONJUNCTION AND DISJUNCTION FALLACIES IN MEMORY

There are two main goals of Experiment 1, and the design reflects these goals. The first goal is to use modeling techniques to identify the memory processes that are the root cause of the discrepancies observed in measures of overdistribution when using indirect versus direct methods. In other words, what processes are different when looking at memory disjunction fallacies and conjunction illusions respectively? To answer this question, the design should have manipulations that affect disjunction and conjunction errors differently. In the prior research, there was evidence for target recollection being important for disjunction fallacies, and context recollection being important for conjunction illusions, so manipulations that would affect target and context recollection variably would be ideal. List order is one such manipulation that has been used in previous studies (Brainerd, Holliday, Nakamura, & Reyna, 2014; Nakamura & Brainerd, 2016). I hypothesize that target recollection will be greater for targets presented in earlier lists, and context recollection will be greater for targets presented in later lists, and that disjunction fallacies will be greater for targets from earlier lists and conjunction fallacies will be greater for targets from later lists. A two-list source monitoring design would be sufficient to test list-order effects for disjunction and conjunction fallacies (Brainerd et al., 2014). This will be the first time that the two recollections will be measured in this paradigm using a model to pinpoint the process differences between the two fallacies.

The second manipulation is categorization. According to the modified dual recollection model, both fallacies depend on phantom context recollection. In prior research, the disjunction

fallacy was caused by an increase in accepting the wrong probe (e.g. accepting L_2? for List 1 targets) rather than a decrease in accepting the disjunctive probe (Nakamura & Brainerd, 2016), which includes the phantom context recollection parameter. The conjunction illusion is entirely dependent on phantom context recollection according to the model equation. Therefore, a manipulation that affects this parameter should see an increase in both fallacies. Because phantom recollection increases with high gist in conjoint recognition, it is reasonable to predict that phantom context recollection would also increase with greater gist, and will be reflected in model parameter estimates (Brainerd et al., 2001). I implemented the categorization manipulation from Nakamura & Brainerd's (2016) study, where some words were categorized on both lists (high gist), while other words were uncategorized (low gist). This also allows for the possibility to look for disjunction and conjunction fallacies for related distractors: if the fallacies occur for distractors, then target recollection is not a necessity for the fallacies to occur. Based on prior data, I expect the conjunction illusions to be present for related distractors, but not disjunction fallacies.

The second goal of Experiment 1 is to test model fit for the new, proposed modified DR model that fits not only disjunction fallacy data, but also conjunction fallacy data. To date, no model exists that accounts for both types of experiments, making it impossible to compare parameter estimates of memory processes from them. In addition, because process comparisons are of key interest, it would be ideal for both disjunction and conjunction fallacies to be tested in the same experiment. This is simple to accommodate: rather than have just three types of probes during test (e.g. L_1?, L_2?, $L_1 \cup L_2$?), there will be all four types of probes during test (e.g. L_1?, L_2?, $L_1 \cup L_2$?, $L_1 \cap L_2$?). The proposed model should fit fine regardless of whether disjunctive and conjunctive probes appear on the same test.

Method

Participants

66 introductory psychology students (48 female, 18 male) were recruited from the same pool used by Brainerd, Holliday et al. 2014. The students participated to fulfill a course requirement.

Materials

The source memory design used in prior research to study memory disjunction and conjunction fallacies was modified for the present experiment. 12 categories were pulled from the Uyeda and Mandler (1980) prototypicality norms. All of the words from this pool are concrete nouns. Of the 12 categories, 8 categories were randomly selected to appear on the study lists, while the remaining 4 categories appeared as unrelated categorized distractors at test. For each of the 8 categories presented at study, 12 words with the highest typicality were chosen to be used for the experiment, and of these, 4 words were presented on List 1, 4 words were presented on List 2, and 4 words were used as related distractors at test. For each of the 4 categories used as unrelated categorized distractors for bias correction data, 8 words with the highest typicality were chosen for the experiment. Finally, 32 concrete nouns that did not appear on any of the lists from the prototypicality norms were chosen from the Toglia & Battig (1978) database, and they were chosen on the basis of equivalent word frequency values as the categorized words.

Two study lists were constructed for the experiment. Both lists comprised of 72 concrete nouns. Each list had 8 categories of 4 words each that were shared across both lists; however, none of the words appeared on both List 1 and List 2. (e.g. the category “bird” appeared on both lists, but List 1 had “eagle,” “hummingbird,” “chicken,” and “starling,” while List 2 had “robin,”

“bluejay,” “swallow,” “oriole”). Words that were categorized always appeared in blocks of four to make the categories salient to participants. In addition to categorized words, each list contained 32 uncategorized words. To prevent participants from adopting a metacognitive strategy to reject all conjunctive probes because they could not remember any target having been presented on both lists, I interspersed the study lists with 4 words that appeared on both lists *but were never tested*. Last, two opening and two closing buffers unrelated to any of the presented words were included on both lists.

As in prior studies of memory disjunction and conjunction fallacies, each study list was accompanied by distinctive visual cues such as font (e.g. Arial on List 1, Stencil on List 2) and background color (e.g. Yellow on List 1, White on List 2). Therefore, each list could be identified by a certain combination of font and background color, bolstering the contextual cues associated to the lists.

The test contained 224 words. Of these, 32 words were targets that had been studied: 8 categorized targets from List 1, 8 categorized targets from List 2, 8 uncategorized targets from List 1, and 8 uncategorized targets from List 2. There were also 32 related distractors from categories that appeared at study but were never presented themselves, 32 unrelated categorized distractors that were from the 4 categories that never appeared at study, and 32 unrelated uncategorized distractors. Each of these cues were factorially paired with one of four possible episodic descriptions: “It was on List 1,” “It was on List 2,” “It was on List 1 or List 2,” “It was on List 1 and List 2.” Participants were asked to accept or reject each probe depending on whether they judged the episodic description to be true or false. Two randomized versions of the experiment were created, and both versions were counterbalanced for a total of four different tests.

Procedure

At the start of the experiment, participants were informed that they would be taking part in a memory test and should try to remember the words presented to the best of their ability. The words were presented at a 3-s rate, centered on the screen and printed in 72-point font. After the first list was completed, there was a 15-s pause before the second list. Once the participants had finished studying the second list, they were provided instructions for the upcoming memory test. Participants were told that during the test, they would see words that they had seen before as well as words that were new, and that they would have to make judgments on whether the episodic descriptions paired with the word were accurate. They were specifically instructed to only accept the test probes if they believed them to be true, and to reject them otherwise. Examples problems were provided to the participants so that they would understand and be familiar with the task. The 224 test words were presented in a random order one at a time, and the participants completed the test in a self-paced manner.

Results

Disjunction Fallacies

The following analysis was conducted on bias corrected (two-high-threshold method; 2HT) acceptance probabilities for both targets and related distractors. Bias corrected values and their calculations are reported on Table 2.1, and raw values of unrelated distractors used to calculate bias correction are reported on Table 2.2. The 2HT correction is a method that has been used in prior research on memory disjunction and conjunction fallacies, and alternative methods of bias correction did not yield significant differences in the results (Brainerd et al., 2012; Brainerd, Holliday, et al., 2014; Brainerd et al., 2016).

Table 2.1
Bias Corrected Acceptance Probabilities

Word content	List-context/statistic			
	$p(L_1)$	$p(L_2)$	$p(L_1 \cup L_2)$	$p(L_1 \cap L_2)$
	Targets			
List 1				
Categorized	.41(.27)	.23(.27)	.46(.29)	.12(.17)
Uncategorized	.35(.22)	.37(.28)	.48(.27)	.08(.16)
List 2				
Categorized	.15(.24)	.43(.30)	.37(.27)	.15(.20)
Uncategorized	.12(.23)	.43(.27)	.40(.29)	.06(.16)
Related Distractors	.00(.21)	.03(.22)	.01(.20)	.01(.12)

Note. Bias was corrected for Categorized (List-specific and Shared) items by subtracting the raw acceptance values of Categorized targets and related distractors by Unrelated Categorized distractors. Bias was corrected for Uncategorized items by subtracting the raw acceptance values of Uncategorized targets by Unrelated Uncategorized distractors. Standard deviations are in the parentheses.

Table 2.2
Unrelated Distractor Acceptance Probabilities

Word content	List-context/statistic			
	$p(L_1)$	$p(L_2)$	$p(L_1 \cup L_2)$	$p(L_1 \cap L_2)$
Distractor Only Category	.24(.21)	.22(.20)	.33(.21)	.06(.11)
Uncategorized	.26(.21)	.17(.16)	.27(.23)	.09(.15)

Note. Standard deviations are provided in the parentheses

The disjunction fallacy occurs when the sum of the probability of accepting "It was on List 1", $p(L_1)$, and the probability of accepting "It was on List 2", $p(L_2)$, is greater than the probability of accepting "It was on List 1 or List 2", $p(L_1 \cup L_2)$. One sample t-tests with a test value of 0 were conducted on the disjunction fallacy index, $p(L_1) + p(L_2) - p(L_1 \cup L_2)$. For targets, disjunction fallacies were observed for all List 1 conditions and they were also observed for all List 2 conditions except for Uncategorized High Frequency targets (Figure 2.1, Table 2.3). As in prior research, disjunction fallacies were not observed for related distractors ($p > 0.6$).

A 2 (list: 1 vs. 2) × 2 (category: categorized vs. uncategorized) × 2 (frequency: high vs. low) ANOVA was conducted on the disjunction fallacy index for targets. There was a main effect of List, $F(1,65) = 15.19$, $p<0.001$, partial $\eta^2 = 0.20$, with higher levels of disjunction fallacy observed on List 1 than on List 2.

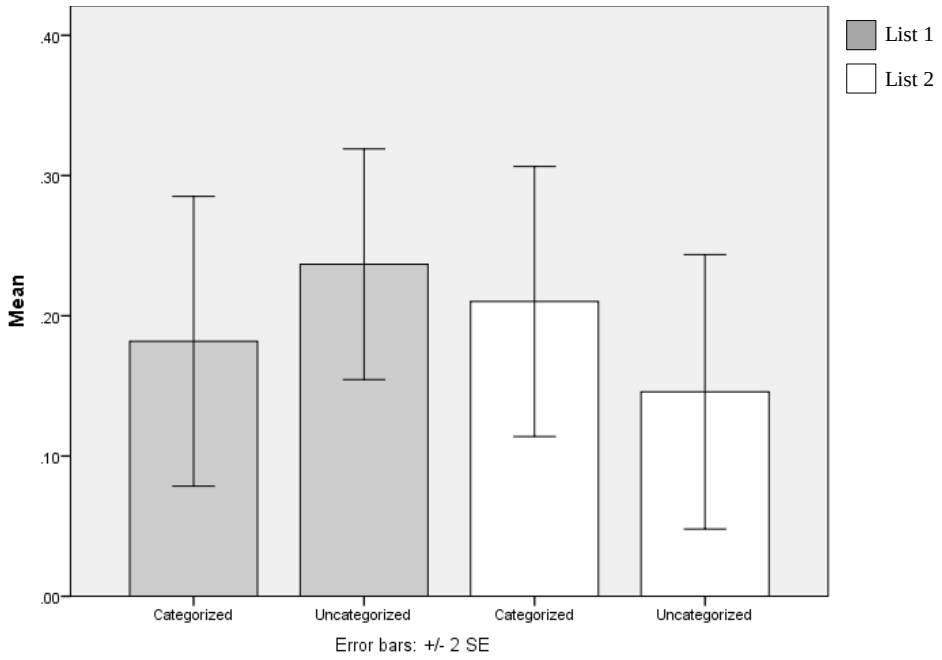


Figure 2.1. Bias corrected disjunction effect for targets.

Table 2.3
Disjunction fallacy index $p(L_1)+p(L_2) – p(L_1\text{-}L_2)$

	mean	*p* value
List 1		
Categorized HF	0.22	0.003
Categorized LF	0.15	0.022
Uncategorized HF	0.15	0.047
Uncategorized LF	0.33	<0.001
List 2		
Categorized HF	0.22	0.003
Categorized LF	0.20	0.002
Uncategorized HF	0.09	0.16
Uncategorized LF	0.20	0.002
Related Distractors HF	0.00	1.00
Related Distractors LF	0.03	0.60

Acceptance Rates

There are two ways a disjunction fallacy may occur: if the acceptance probabilities of the nondisjunctive probes are high, or if the acceptance probability of the disjunctive probe is low. A 2 (category) × 2 (list) × 2 (frequency) × 4 (probe: accept correct vs. accept incorrect vs. accept disjunction vs. accept conjunction) was conducted on bias corrected acceptance probabilities for targets. Because list effects were observed for disjunction fallacies, there was a particular interest in seeing if there was a two-way List × Probe interaction effect. There was, $F(3, 195) = 14.58$, partial $\eta^2 = 0.18$. According to post-hoc comparisons, the incorrect question was accepted more often for List 1 targets ($L_2?|L_1$) than for List 2 targets ($L_1?|L_2$) (mean difference = 0.17, SE = 0.03, $p<0.001$). The disjunctive question was also accepted more for List 1 targets than for List 2 targets (mean difference = 0.09, SE = 0.02, $p<0.001$). The observation that disjunction fallacies were greater for List 1 than for List 2 is, based on these results, most likely due to the greater acceptance of the incorrect probe for List 1 targets. This result was also found in prior research (Nakamura & Brainerd, 2016).

Conjunction Illusions

The following analysis was conducted on bias corrected (2HT) acceptance probabilities for both targets and related distractors. Bias corrected values and their calculations are reported on Table 2.1, and raw values of unrelated distractors used to calculate bias correction are reported on Table 2.2. Conjunction illusions occur when the probability of accepting the conjunctive probe, $p(L_1 \cap L_2)$, is greater than zero. Conjunction fallacies are stronger iterations of the conjunction illusion where not only is $p(L_1 \cap L_2)$ greater than zero, which is illogical, but it is greater than one or the other of the nonconjunctive probes.

To determine if there were any conjunction illusions, a one sample t-test was conducted on $p(L_1 \cap L_2)$. Conjunction illusions were observed on all conditions for both List 1 and List 2 for targets, and they were also observed for high frequency related distractors (Figure 2.2, Table 2.4). A 2 (list: 1 vs. 2) × 2 (category: categorized vs. uncategorized) × 2 (frequency: high vs. low) ANOVA was conducted on the conjunction illusion index for targets. There was a main effect of Category, $F(1,65) = 8.06$, $p=0.006$, partial $\eta^2 = 0.11$, where conjunction illusions were elevated for categorized targets compared to uncategorized targets.

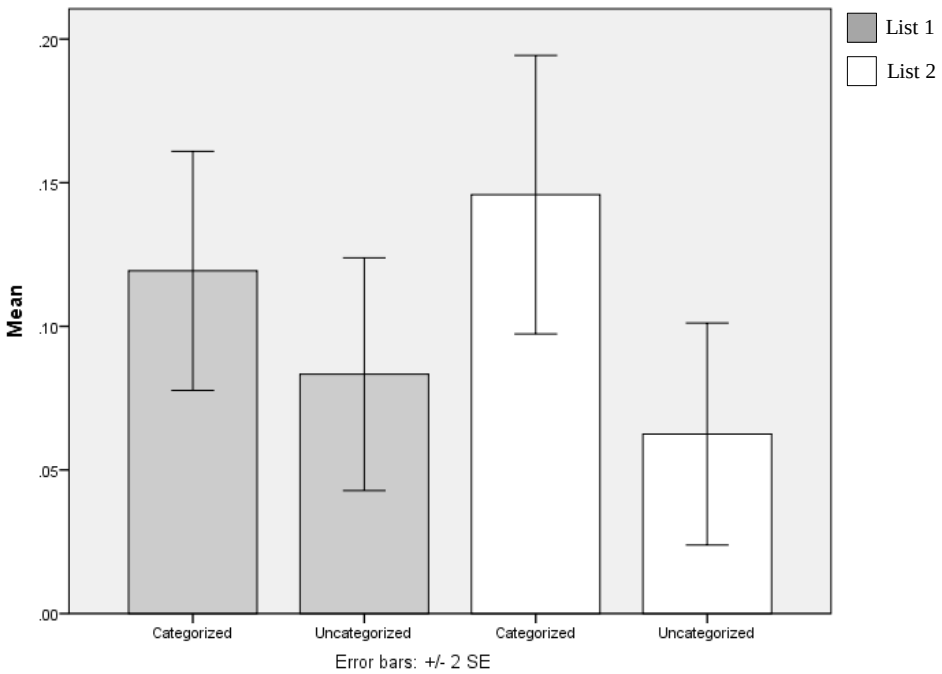


Figure 2.2. Bias corrected conjunction effect for targets.

Table 2.4
Conjunction illusion index $p(L_1 \cap L_2)$

	mean	*p* value
List 1		
Categorized HF	0.12	<0.001
Categorized LF	0.12	<0.001
Uncategorized HF	0.09	0.002
Uncategorized LF	0.08	0.004
List 2		
Categorized HF	0.18	<0.001
Categorized LF	0.11	0.001
Uncategorized HF	0.07	0.03
Uncategorized LF	0.06	0.02
Related Distractors HF	0.05	0.02
Related Distractors LF	-0.03	0.07

To determine if any of the conjunction illusions became conjunction fallacies, a paired t-test was conducted on the probability of accepting the conjunctive probe, $p(L_1 \cap L_2)$, and the probability of accepting either of the nonconjunctive probes, $p(L_1)$ and $p(L_2)$. Conjunction fallacies were not observed in this experiment.

Model Results

Of key interest to this study was whether a revised dual recollection model could produce acceptable fits of the data from both disjunction and conjunction designs. Because there is evidence that different memory processes operate on disjunction and conjunction fallacies, I was interested in whether model results would support this prior finding (Nakamura & Brainerd, 2016).

The modified dual recollection model contains seven parameters (C, P, F, T, b, b_{1U2}, *and* $b_{1 \cap 2}$), and the present design provides eight free probabilities with which to estimate

them. Therefore, the fit test for this model to determine if it is acceptable to account for the response probabilities is a $G^2(1)$ statistic, with a critical value 3.84 to reject the null hypothesis of fit at the 0.05 level. The null hypothesis could not be rejected for all conditions except for high frequency uncategorized items for both List 1 and List 2 (Table 2.5).

Table 2.5
Model fit results for Revised Dual Recollection Model

	G^2	*C*	*T*	*F*	*P*	*b*	b_{12}	$b_{1\cap 2}$
L_1 Categorized HF	3.07	0.24	0.41	0.13	0.54	0.27	0.41	0.05
L_1 Categorized LT	0.47	0.19	0.43	0.47	0.70	0.18	0.25	0.08
L_1 Uncategorized HF	11.25	0.09	0.37	0.26	1.00	0.28	0.34	0.13
L_1 Uncategorized LF	0.47	0.19	0.43	0.47	0.70	0.18	0.25	0.08
L_2 Categorized HF	2.05	0.20	0.38	0.00	0.94	0.27	0.42	0.05
L_2 Categorized LT	0.50	0.32	0.38	0.00	0.40	0.18	0.24	0.08
L_2 Uncategorized HF	10.93	0.14	0.27	0.13	0.56	0.28	0.34	0.12
L_2 Uncategorized LF	3.30	0.30	0.41	0.06	0.20	0.15	0.20	0.05

The results of the model fits were surprising, considering how the dual recollection model has fit very well for uncategorized items in prior studies (e.g. Brainerd et al., 2012). In order to see if the poor fits were due to the modifications to the original dual recollection model, fit tests were run using the prior DR model using only the data from the disjunction fallacy part of the design. With five parameters with six free empirical probabilities to estimate them, the fit test is a $G^2(1)$ statistic with a critical value of 3.84 to reject the null hypothesis of fit at the 0.05 level. Again, only high frequency uncategorized items for both List 1 and List 2 failed the test.

One possibility for poor model fits, especially with regards to the two high frequency uncategorized conditions, is that assumptions made about the data were not necessarily true. Namely, the model assumes that there is an absence of individual differences in parameter

values, but this may not have been the case (Brainerd et al., 2015). For this study I have followed the conventional approach to modeling; that is, to first identify if a model fits aggregated data and then to move on to models that allow for individual differences. The relevant test for the present study is a X^2 statistic of the following form (Smith & Batchelder, 2008):

$$X^2(N-1) = \sum_{i=1}^{N}\{(R_i - R_e)^2 / R_e\} + (M - R_i - R_e^*)^2 / R_e^*\} \qquad (15)$$

In this test, N is the number of subjects, R_i is the total number of acceptances of a particular cue-probe combination (e.g., L_1?|HFL_1targets), R_e is $(\sum_{i=1}^{N} R_i)/ N$, $R_e^* = M - R_e$, and M is the total number of the particular cue-probe combination that subjects responded to. There were four unique target cue-probe combinations and four distractor cue-probe combinations for the two experimental conditions (e.g. List 1 uncategorized high frequency words) whose data did not provide good fits to the model; therefore, sixteen tests for individual differences were conducted. Because $N = 66$, the critical value of X^2 to reject the null hypothesis of no individual differences at the 0.05 level was 84.82.

Out of the sixteen tests, for both L_1 and L_2 high frequency uncategorized items, the null hypothesis was rejected at a high level of confidence for disjunctive ($L_1 \cup L_2$?) and conjunctive ($L_1 \cap L_2$?) probes for distractors in both L_1 (disjunctive: 96.88, conjunctive: 104.41) and L_2 (disjunctive: 96.88, conjunctive: 104.41) experimental conditions. Considering how the model's predicted values for the bias parameters varied greatly with observed values though not as much for other parameters, it appears that model fit failed because of individual differences in response bias. The presence of individual differences in the data suggests that while the aggregated data did not produce acceptable fits, individual data may fit. The next step therefore was to fit the model to individual data for both L_1 and L_2 high frequency uncategorized conditions. If the

model is correct, the values for the $G^2(1)$ statistic for individualized fits will be within acceptable boundaries when the fit values are averaged across the subject sample. The mean G^2 statistics for the two conditions were 3.26 and 3.11 for L_1 and L_2 respectively, indicating that when individual differences were accounted for, the model provided acceptable fits. In addition, 67% of the participants had $G^2(1) < 3.84$ for the L_1 high frequency uncategorized condition, and 70% of the participants had $G^2(1) < 3.84$ for the L_2 high frequency uncategorized condition. Therefore, we can attribute fit failures for the aggregated data in these two conditions to the violation of the assumption of no individual differences. In sum, the revised dual recollection model provides acceptable fits for all 8 experimental conditions.

Parameter Comparisons

Estimates for the parameter values for the revised dual recollection model are reported on Table 3.5. The main purpose for the following analysis is to determine if there are any differences in memory processes depending on condition differences such as the list the target was presented on, or its relative frequency. We can determine this using a series of likelihood ratio tests. First we run an omnibus test with a null hypothesis $CatL_1HF = CatL_1LF = CatL_2HF = CatL_2LF = UncatL_1HF = UncatL_1LF = UncatL_2HF = UncatL_2LF$, with a $G^2(7)$ statistic with a critical value of 14.07, for each of the four memory parameters (C, P, F, T), and three bias parameters (b, b_{1U2}, $b_{1\cap 2}$). If the null hypothesis is rejected, then we run likelihood ratio tests separated by category ($L_1HF = L_1LF = L_2HF = L_2LF$ for categorized items, and $L_1HF = L_1LF = L_2HF = L_2LF$ for uncategorized items) and list ($CatHF = CatLF = UncatHF = UncatLF$ for List 1 items, $CatHF = CatLF = UncatHF = UncatLF$ for List 2 items) with a $G^2(3)$ statistic with a critical value of 7.82. Following the rejection of the null hypothesis for those sets of analyses,

we move on to pairwise tests with a $G^2(1)$ statistic with a critical value of 3.84 to identify any parameter differences based on Categorization, List, and Frequency.

The first omnibus test produced null hypothesis rejection for six of the seven parameters: context recollection (16.28), phantom context recollection (16.90), familiarity (23.79), *b* (71.51), b_{1U2} (60.96), and $b_{1\cap 2}$ (19.60). For the next set of omnibus tests separated within categories, familiarity (13.23), *b* (29.84), and b_{1U2} (35.16) produced null hypothesis rejections for categorized items, and familiarity (9.56), *b* (41.53), b_{1U2} (20.89), $b_{1\cap 2}$ (11.41) produced null hypothesis rejections for uncategorized items. For tests separated within lists, *b* (30.42), b_{1U2} (23.47), and $b_{1\cap 2}$ (8.36) produced null hypothesis rejections for List 1 items, and phantom context recollection (14.29), *b* (40.90), b_{1U2} (37.10), and $b_{1\cap 2}$ (11.00) produced null hypothesis rejections for List 2 items. A total of 56 pairwise tests were conducted based on the results of the omnibus tests (12 categorized, 16 uncategorized, 12 list 1, 16 list 2). I will discuss the key findings from the tests for each parameter in the following section, first going over the memory parameters and then the bias parameters.

For phantom context recollection, *P*, there was a frequency effect where List 2 categorized items had greater *P* for high frequency items than for low frequency items (*LF=HF*, $G^2(1)=5.40$). For familiarity, frequency effects were observed for categorized items on List 1 only, where familiarity was greater for low frequency items than high frequency items (*LF=HF*, $G^2(1)=5.87$). List effects were also observed with familiarity. Familiarity was greater for L_1 than for L_2 for low frequency items for both categorized ($L_1=L_2$, $G^2(1)=10.85$) and uncategorized ($L_1=L_2$, $G^2(1)=8.12$) items. No pairwise comparisons were done for target recollection or context recollection as there were no null hypothesis rejections for the omnibus tests on *T* or *C*.

The model contains three bias parameters, and all three had similar patterns with regards to experimental manipulations. Bias *b* was greater for high frequency items than low frequency items for both categorized List 1 (*LF=HF*, $G^2(1)=14.67$) and List 2 (*LF=HF*, $G^2(1)=15.16$), and uncategorized List 1 (*LF=HF*, $G^2(1)=15.74$) and List 2 (*LF=HF*, $G^2(1)=25.33$). The same frequency pattern was observed for b_{1U2}, with high frequency having greater b_{1U2} than low frequency items for categorized List 1 (*LF=HF*, $G^2(1)=16.02$) and List 2 (*LF=HF*, $G^2(1)=19.13$), and uncategorized List 1 (*LF=HF*, $G^2(1)=6.17$) and List 2 (*LF=HF*, $G^2(1)=13.97$). Last with $b_{1\cap 2}$, it was greater for high frequency items than low frequency items for uncategorized List 2 words (*LF=HF*, $G^2(1)=7.91$). Bias for conjunctive probes also had a category effect that was not observed for the other two kinds of bias, where it was greater for high frequency uncategorized items than for categorized items on List 1 (*Cat=Uncat*, $G^2(1)=7.91$) and List 2 (*Cat=Uncat*, $G^2(1)=7.91$

Based on the initial hypotheses that target recollection has list effects, context recollection has list and categorization effects, phantom recollection has categorization effects, and familiarity has frequency effects, it is worth checking if the values of these parameters are in the predicted direction, even if significant differences were not observed at the pairwise level of the omnibus tests. According to the omnibus tests, target recollection was not significantly different across all of the experimental manipulations. However, the direction is still consistent with the prediction that it is greater on List 1 than on List 2 ($L_1=0.29$; $L_2=0.24$). Regarding *C*, there were no significant differences observed at the pairwise level, but there was a main effect. Again, the direction of differences was in the predicted level, with *C* being greater on List 2 than on List 1 ($L_1=0.32$; $L_2=0.36$), and greater for categorized than uncategorized items (Cat=0.37; Uncat=0.31). *P* was predicted to be greater on categorized items than uncategorized items, and

this was supported in the omnibus tests (Cat=0.39; Uncat=0.32). Last, *F* was predicted to be greater for low frequency items than for high frequency items, and this was also supported in the omnibus tests (HF=0.14; LF=0.25). Unexpectedly, *F* was greater for List 1 words than for List 2 words as well (L_1=0.34; L_2=0.05).

Discussion

There were two goals in Experiment 1. The first was to see whether indirect and direct measures of overdistribution were equitable, or if their indexes, disjunction fallacies and conjunction illusions respectively, tapped different memory processes. The second was to propose a revised dual recollection model to accommodate memory conjunction illusion designs, so that memory processes affecting disjunction and conjunction fallacies could be directly compared. Without the model, it was not possible to run direct process comparisons, and prior experiments relied on inferences.

With regards to the first goal, results replicated findings by Nakamura & Brainerd (2016) where the disjunction and conjunction fallacies were affected differently by certain manipulations. Results fit the prediction that target recollection was necessary for disjunction fallacies, whereas it was not for conjunction illusions: while disjunction fallacies were only present for targets, conjunction illusions were observed for both targets and related distractors. In addition, while list order effects were present for disjunction fallacies with the fallacy being greater for List 1 than List 2 as predicted, there were no such list order effects for conjunction illusions. Instead, categorization was the determining factor for conjunction illusions, such that categorized items had greater illusions than uncategorized items. This fits with the prediction that greater gist would increase overdistribution errors, and therefore the conjunction illusion. In addition, there has been prior work showing greater gist supports retrieval of contextual details

for both targets and related distractors (Ball et al., 2014; Chen, Brainerd, & Gomes, in press), and this is consistent with prior studies suggesting that the illusion is driven by context recollection.

What do these differences mean in terms of measuring overdistribution? Recall that overdistribution is when memory traces consistent with multiple memory states can be retrieved, or 'overdistributed,' such that a subject will accept both true and false memories of an event. This definition is embedded in the revised dual recollection model for probability of accepting the conjunctive probe, or rather, the direct measure of overdistribution $p(L_1 \cap L_2)$. In the revised model, $p(L_1 \cap L_2)$ will be accepted when a subject retrieves not only the contextual cues consistent with the true event, *C*, but also when they retrieve contextual cues consistent with the false event, *P*. On the other hand, while the definition of overdistribution (that traces consistent with both the true and false event can be retrieved) is also embedded in the equation for the disjunction fallacy as well, its interpretation is different. The conjunction illusion occurs when *contextual* cues consistent with *both* true and false events are retrieved, but the disjunction fallacy occurs when *target* recollection is available without context recollection. Therefore, the direct measure of overdistribution, the conjunction illusion, is perhaps something that is more consistent with the original proposed definition of overdistribution for source memory designs.

The second goal of the experiment was to compare process differences across conditions to identify what memory processes contribute to disjunction and conjunction fallacies. The revised dual recollection model with the new phantom context recollection parameter provided acceptable fits to the data. According to the model, it was predicted that phantom context recollection was crucial for conjunction illusions (though not necessary for disjunction fallacies), and therefore we should observe parallel increases in *P* where conjunction illusions were elevated. There was one manipulation that affected conjunction illusions: categorized items had

increased conjunction illusions than uncategorized items. If it were true that *P* was important for the illusions, then they should also increase when words were categorized. They did. *P* was greater when words were categorized rather than uncategorized, supporting the hypothesis that it is an important component for conjunction illusions.

One surprising finding in this experiment is the fact that there was no main effect of List for target recollection, unlike what has been observed in prior studies involving uncategorized lists using the same kind of source monitoring design (e.g. Brainerd, Gomes, & Nakamura, 2015; Nakamura & Brainerd 2016). However, the direction of the effect was still supported in this study, with the raw value of *T* being larger for List 1 than for List 2.

It is also possible that the previous assumption that target recollection was the main driving force of disjunction fallacies was not the full story. Few studies have used models to measure process estimates in disjunction fallacy source memory designs. Nakamura & Brainerd (2016) concluded that target recollection may be the driving force of disjunction fallacies based on List order effects, and the observation that the fallacies only occurred for targets. However, they did not use model analyses, and their data can also be supported by a familiarity argument which is also consistent with what was found in the study by Brainerd et al. (2012). Recall that the 2012 study used the CPD model, which does not distinguish between the two recollections, but does separate familiarity. The 2012 study found that the disjunction fallacy increased when *F* increased, which is consistent with the results of the present experiment. This is not to say that target recollection is not important at all: it was still the case that disjunction fallacies only appeared for targets, and not related distractors. The interpretation may be that target recollection is necessary for the fallacy, but the fallacy increases with familiarity as well and the list effects for the present experiment reflect a familiarity effect rather than a target recollection

one. Indeed, Brainerd et al. (2015) analyzed DR model parameter estimates using a categorized source memory design, and found list effects for target recollection that wasn't observed in the present experiment.

Last, list effects were predicted for conjunction illusions in line with the findings from Nakamura & Brainerd (2016), but no such effects were observed in the present experiment. However, this does not mean that the hypothesis that context recollection supports conjunction illusions is incorrect. First, while there were no statistically significant differences in context recollection across lists, the direction was still consistent with predictions (L_1=0.32, L_2=0.36). Had the list differences in context recollection been significant, as was found in Brainerd et al.'s (2015) study, there may have been statistically significant list effects for the conjunction illusion. Second, previous experiments did not include the phantom context recollection parameter, which not only was the parameter that was defined as a necessary process for conjunction illusions to occur in the model, but was also predicted to be sensitive to gist manipulations. The sensitivity of conjunction illusions to gist manipulations is consistent with findings from prior studies, and fits the predictions made by the revised DR model.

To summarize, the revised DR model provided new evidence on the nature of disjunction and conjunction fallacies in memory. While both types of memory fallacies were originally proposed to measure overdistribution errors indirectly and directly, there is growing support that indirect measures of overdistribution may be measuring something else. While the 'direct' measure, conjunction illusions, are affected primarily by phantom context recollection, the 'indirect' measure, disjunction fallacies, seem to be affected by target recollection and familiarity.

CHAPTER 3

INTRODUCTION: LATENCY EXTENSION MODELING

As Experiment 1 demonstrated, models allow us to test hypotheses about phenomena directly, rather than relying on inferences. In the case of the first experiment, the revised dual recollection model provided insight into how target recollection and familiarity drove disjunction fallacies and phantom context recollection drove conjunction illusions, and why calling the two fallacies 'indirect' and 'direct' measures of overdistribution may not be entirely accurate. For the remaining experiments of this thesis, I will discuss how a new model is able to provide estimates of something that could not previously be measured: relative process speeds.

A new methodology proposed by Brainerd, Nakamura, and Lee (in press) opened the possibility to allow models to directly measure the relative speed of (for example) memory processes. For Experiment 2, I implemented this latency extension procedure to an existing memory model to show how it can be applied to real data, as well as answer questions about process speeds that could not be answered previously. While Experiment 2 used preexisting data to evaluate the fit of latency extension models and gather preliminary data on process speeds, Experiment 3 was designed specifically to test follow-up questions that emerged from Experiment 2. I will first discuss the general hypotheses of memory process speeds in the literature, talk about the model that was used in Experiment 2 and Experiment 3 to answer some of these questions, and then discuss how this new latency extension model compares to other contemporary models that incorporate time course of memory processes.

With a greater understanding of the workings of memory, one question that remains unclear is whether different memory processes become available at different speeds. In the early 1970s, it was proposed that in recognition, vague feelings of having seen an item before, or

familiarity, was faster than retrieving vivid details of presented targets, or recollection (Atkinson & Juola, 1973, 1974). It was thus suggested that fast recognition decisions were mostly due to fast familiarity, and slower decisions were due to recollection. Mandler (1980) formalized this with a theoretical explanation on why familiarity and recollection operated at different speeds: recollection was only activated if familiarity failed to provide enough evidence to make a recognition decision. Matching with theories that propose recollection to be a more effortful process than familiarity, it seems plausible that the more cognitively demanding process would be slower than the less demanding one (Gardiner, 1988; Jacoby, Toth, Yonelinas, & Debner, 1997).

However, studies that tested the hypothesis of familiarity being faster than recollection did not provide conclusive evidence for it. Gillund and Shiffrin (1984) tested whether familiarity is faster than recollection by using a speeded response recognition design. In their experiment, they had subjects make standard old/new recognition judgments to different types of cues (targets and distractors) in fast and slow timed conditions. In fast conditions, subjects were required to make a decision within 900ms of a cue's presentation, while in slow conditions, subjects were not allowed to respond until 1s had elapsed since a cue's presentation. The experiment contained different manipulations that should affect recollection, such as list length, number of presentations, and depth of encoding. If it were true that recollection was a slow process, then manipulations that affect recollection should have a greater effect at longer decision times. This was not observed in their study.

In addition, remember-know (RK: Tulving, 1985) studies fail to support the hypothesis that familiarity is faster than recollection. The RK paradigm is a recognition task that requires subjects to identify how they based their memory decision. When subjects indicate that a cue is

"old," they can make either a "remember" (R) or a "know" (K) response. R indicates that the recognition judgment was accompanied by conscious recollection of details of the item during presentation (mapping onto recollection), while K indicates that a recognition judgment was not accompanied by such details (mapping onto familiarity). If it is the case that familiarity is faster than recollection, we might expect K responses to be faster than R. However, studies have shown the opposite to be true: R responses are faster than K (Dewhurst & Conway, 1994; Dewhurst, Holmes, Brandt, & Dean, 2006).

Is it the case, then, that the hypothesis that familiarity is faster than recollection is incorrect? There is no definitive conclusion for this, either, for some results observed in the false memory literature provide support for it. In particular, the phenomenon of the inverted U shape of false memory over time can be explained by the idea that familiarity is fast and recollection is slow. Brainerd, Reyna, Wright, & Mojardin (2003) proposed that the inverted U for false memory occurs because the process that supports false memory according to FTT (familiarity) is fast while the process that supports true memory (recollection) is slow. They reported findings from studies where subjects studied a list of semantically related words and were later asked to make old/new decisions on targets (e.g. spruce, oak), related distractors (e.g. pine, maple), and unrelated distractors (e.g. car, bagpipe). The old/new decisions were made after various time intervals, a standard response-signal procedure (Dosher, 1984). A signal to respond was made to subjects after some lag (e.g. 250, 500, 750, 1,000, and 1,500msec) following probe presentation, and they were asked to make a decision within a brief interval following the signal (e.g. 300msec). Indeed, false memory to related distractors increased over time before decreasing to form an inverted U in this study. Brainerd et al. (2003) argued that as fast familiarity initially increased without slow recollection, false alarm rates to related distractors increased, forming the

left arm of the curve. Then, as recollection became available, false alarms decreased over time, forming the right arm of the curve.

The weakness in all of the designs that have been discussed thus far is that they are, ultimately, indirect measures of process speeds and rely heavily on inference. In addition, all discussion on familiarity versus recollection assumes that recollection is a univariate process. However, as described in Experiment 1, dual-recollection theory introduces a bivariate model with two types of recollection: target and context recollection (Brainerd, Gomes, et al. 2014). In semantic false memory designs, target recollection and context recollection supports true and false memory respectively, and it is unclear whether these two recollections have different time courses as well. It could very well be the case that fast context recollection is responsible for the left arm of the inverted U curve of false memory, and that a mix of familiarity and target recollection is responsible for the right arm of the curve. Without a way to directly measure process speeds, however, it is impossible with current inferential methods to determine if familiarity is faster than both forms of recollection, or whether context recollection is, in fact, faster than both familiarity and target recollection.

Conjoint Recognition Latency Extension Model

Latency extension models are able to answer the inverted U false memory phenomenon because they are able to directly measure the relative speeds of processes. An appropriate model that can address the familiarity-recollection question needs two important characteristics. First, it must be able to provide estimates for the processes in question: familiarity and the two recollections, target and context recollection. Second, it must be able to estimate the relative speeds of each of these processes.

The conjoint recognition design for semantic false memory addresses both points. As

described in Experiment 1, the conjoint recognition design has subjects study lists of semantically related words, and then take a memory test on targets, related distractors, and unrelated distractors that are paired with one of three possible probes. The probes correspond with a verbatim retrieval condition (V: Was it a target?), a gist retrieval condition (G: Was it a related distractor?), and a verbatim and gist retrieval condition (VG: Was it a target or a related distractor?). The conjoint recognition model itself has parameters associated with targets, related distractors, and unrelated distractors, described in more detail on Table 3.1. There are three retrieval parameters for targets: identity judgment (*I*: a mix of both target and context recollection), erroneous recollection rejection (*E*: context recollection), and familiarity (S_1). There are three retrieval parameters for related distractors: recollection rejection (*R*: target recollection), phantom recollection (*P:* context recollection), and familiarity (S_2). Finally, there are three bias parameters associated with unrelated distractors: bias for V probes (b_V), bias for G probes (b_G), and bias for VG probes (b_{VG}). For the purposes of testing relative speeds of familiarity and recollection, the conjoint recognition model clears the first criteria because it provides separate estimates of the two recollections and familiarity.

Table 3.1

Conjoint Recognition Model parameter definitions

Parameter	Definition
Targets	
I	Identity Judgment: A list item encourages the conscious reinstatement of its prior presentation, or the conscious reinstatement of the context it appeared in during encoding. Identity judgment promotes the acceptance of the item as having been a target, and thus supports its acceptance on V and VG probes, but not for G probes (target and context recollection).
E	Erroneous Recollection Rejection: A list item encourages gist-cued retrieval of verbatim traces of a *different* target, which results in the mistaken rejection of the target probe. This results in its rejection on V probes, and acceptance on G and VG probes (context recollection).

S_1	Semantic Familiarity: A list item's meaning provokes feelings of familiarity that are strong enough to allow it to be perceived as either a target or a related distractor. Items that provoke familiarity but not identity judgment are accepted on V, G, and VG probes (familiarity).
Related distractors	
R	Recollection rejection: A related distractor (e.g. *doctor*) provokes the conscious reinstatement of corresponding targets (e.g. *nurse*), supporting its identification as a distractor and not a target. Recollection rejection encourages distractors to be accepted on the G and VG probes, but not on V probes (target recollection)
P	Phantom recollection: A related distractor encourages the conscious reinstatement of the contextual details that accompanied corresponding targets. Items that provoke the recollection of these contextual details are judged to be targets, thereby supporting its acceptance on V and VG probes, but not on G probes (context recollection)
S_2	Semantic Familiarity: A related distractor's meaning provokes feelings of familiarity that are strong enough to allow it to be perceived as either a target or a related distractor. Familiarity encourages acceptance of related distractors on V, G, and VG probes (familiarity)
Unrelated distractors	
b_V	Bias: high-threshold bias parameter for V probes
b_G	Bias: high-threshold bias parameter for G probes
b_{VG}	Bias: high-threshold bias parameter for VG probes

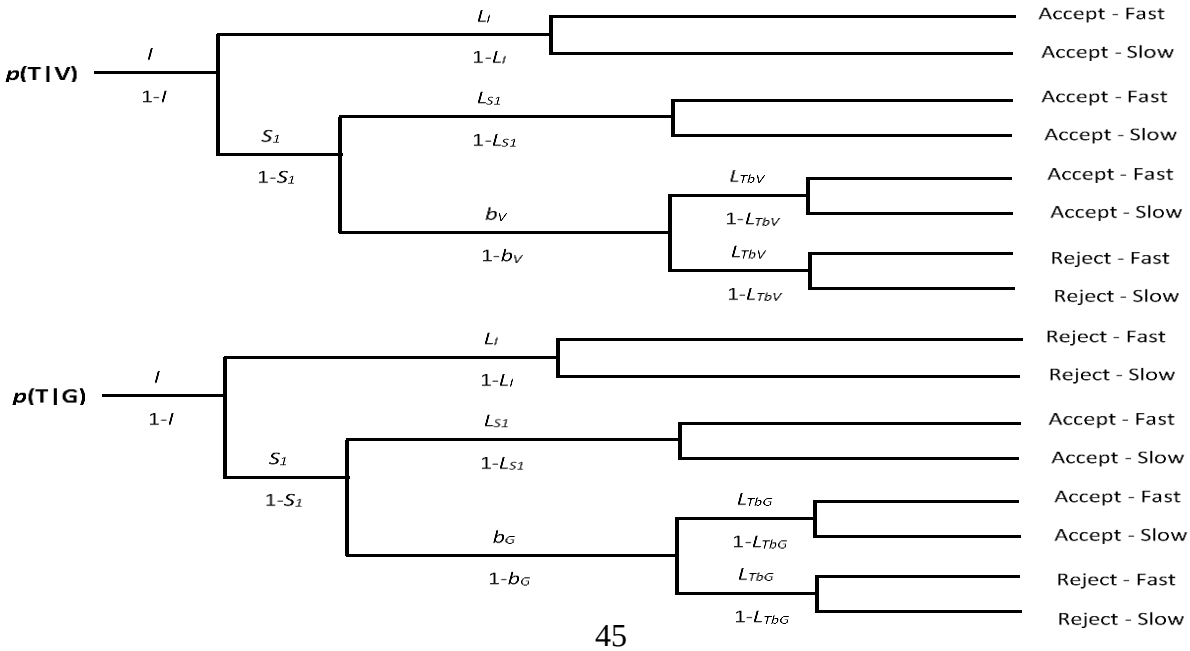
The second criteria can be cleared with a latency extension of the conjoint recognition model. Methods on how to extend a multinomial model to include response latencies are discussed in detail by Heck and Erdfelder (2016, 2017). The general idea behind latency extensions is that different cognitive responses that contribute to a discrete response are composed of mixtures of latency distributions attributed to each process. When responses are binned as fast or slow (e.g. fast acceptance of a probe versus slow accept), the latency

distributions of the cognitive processes can be parameterized in a similar manner that the processes in the core model are parameterized. In summary, the same mathematical procedure that is used to estimate parameters in the core model can be used to estimate relative speeds of those processes in the latency extended conjoint recognition model.

Extending the conjoint recognition model with two speeds of "fast" and "slow" yields a model with thirty-six latent reaction time distributions. In order to make the model identifiable, some theoretically motivated assumptions were made with regards to the values of the latency parameters. The first assumption is that the speed for accepting or rejecting a target or a related distractor based on bias is the same. Second, examining the process trees for the V, G, and VG conditions for accepting a probe based on identity judgment, one can see that the tree structure is identical. Therefore, the speed in which identity judgment becomes available must be assumed equal across the three conditions for the model to be identifiable. The same is true for the other parameters, so that we are left with five latency parameters each associated with a retrieval process (L_I, L_{S1}, L_R, L_P, L_{S2}), six bias latency parameters associated with list items and related distractors (L_{TbV}, L_{TbG}, L_{TbVG}, L_{RDbV}, L_{RDbG}, L_{RDbVG}), three bias latency parameters associated with accepting unrelated distractors (L_{bV}, L_{bG}, L_{bVG}), and three bias latency parameters associated with rejecting unrelated distractors ($L_{(1-bV)}$, $L_{(1-bG)}$, $L_{(1-bVG)}$).

The full tree with all of the assumptions are available on Figure 3.1 (list items), Figure 3.2 (related distractors), and Figure 3.3 (unrelated distractors). In all three figures, the far left describes the probabilities of accepting cues with one of three probes (V, G, and VG). The far right represents one of four possible response outcomes (fast-accept, slow-accept, fast-reject, slow-reject). For the purposes of answering the question on whether recollection or familiarity is faster than the other, we turn to two possible comparisons. The first are latency parameters from

targets, L_I and L_{S1}. Recall that I, identity judgment, is a mixture of the two recollections, and S_1 is familiarity. Comparing the two latency parameters for identity judgment and familiarity can be a test on whether or not the univariate recollection theory holds in its prediction that familiarity is faster than recollection. The second comparison of interest, and the more interesting of the two, are between the latency parameters from related distractors, L_R, L_P, and L_{S2}. Recollection rejection R is target recollection, phantom recollection P is context recollection, and S_2 is familiarity. Comparing the relative speeds of these parameters provides a test to see if fast context recollection, P, is responsible for the left arm of the inverted U curve of false memory, and also whether or not there are any speed differences between target recollection and familiarity. According to univariate theories of recollection, we would expect to see familiarity being faster than both forms of recollection in the model. A bivariate recollection account is supported if context recollection is faster than familiarity, and context and target recollection have different speeds.



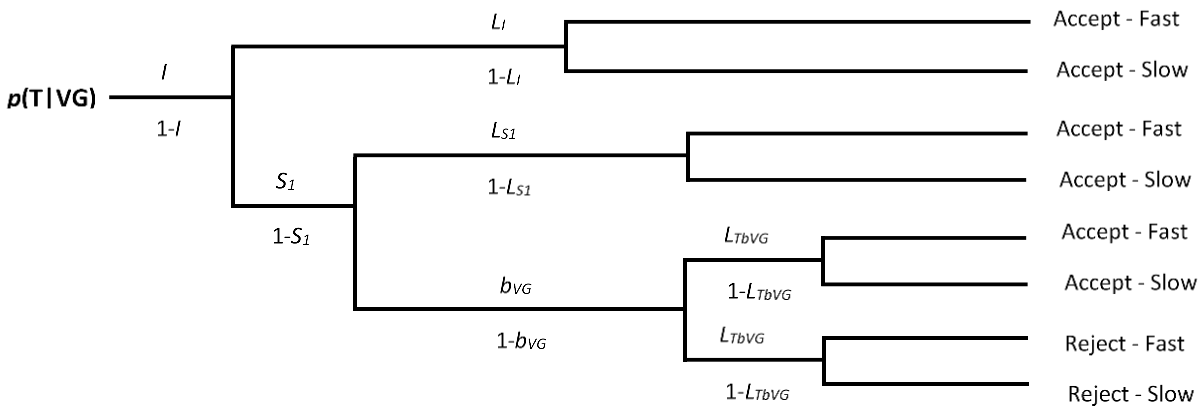
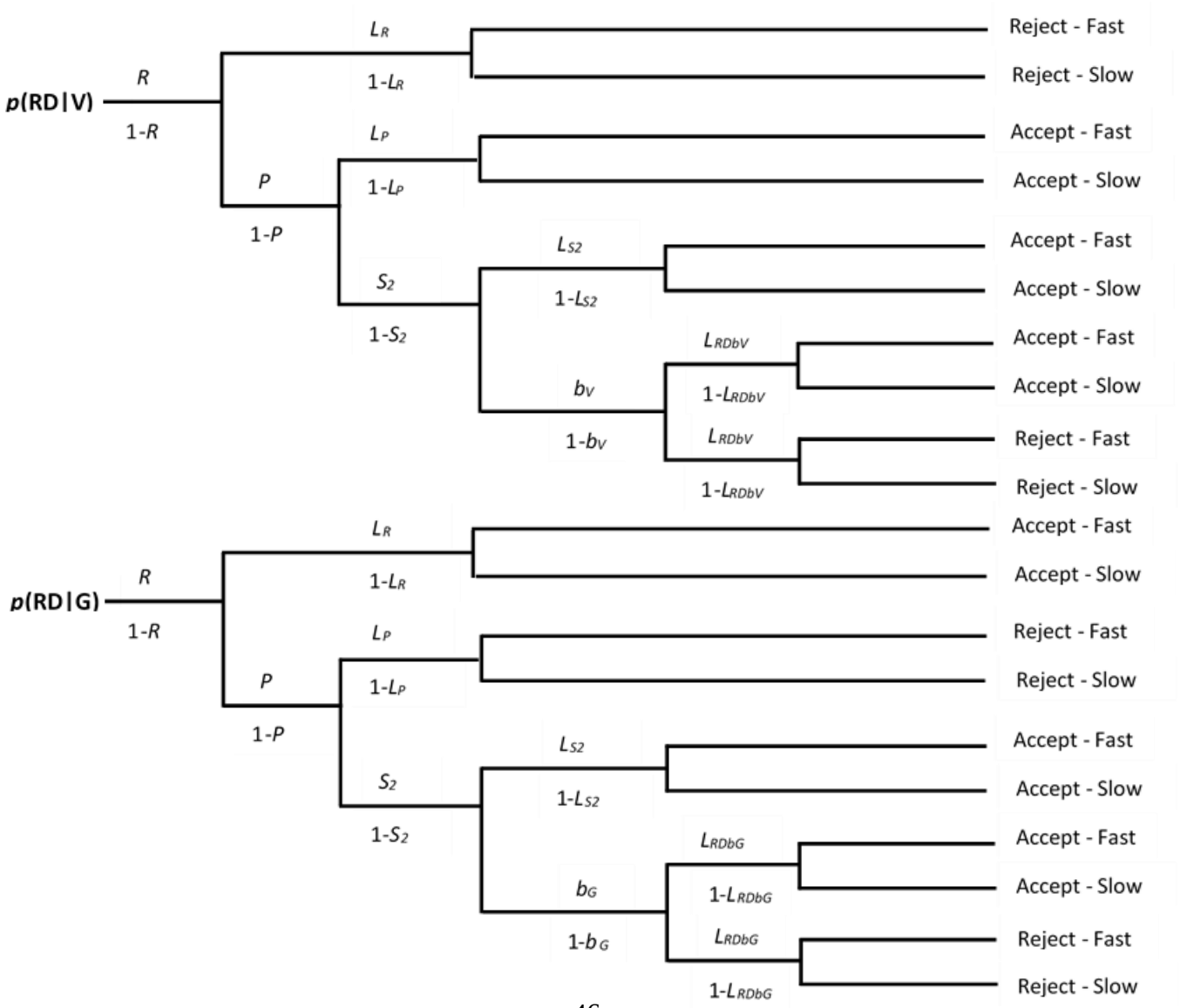


Figure 3.1. Latency extended conjoint recognition model trees for targets.



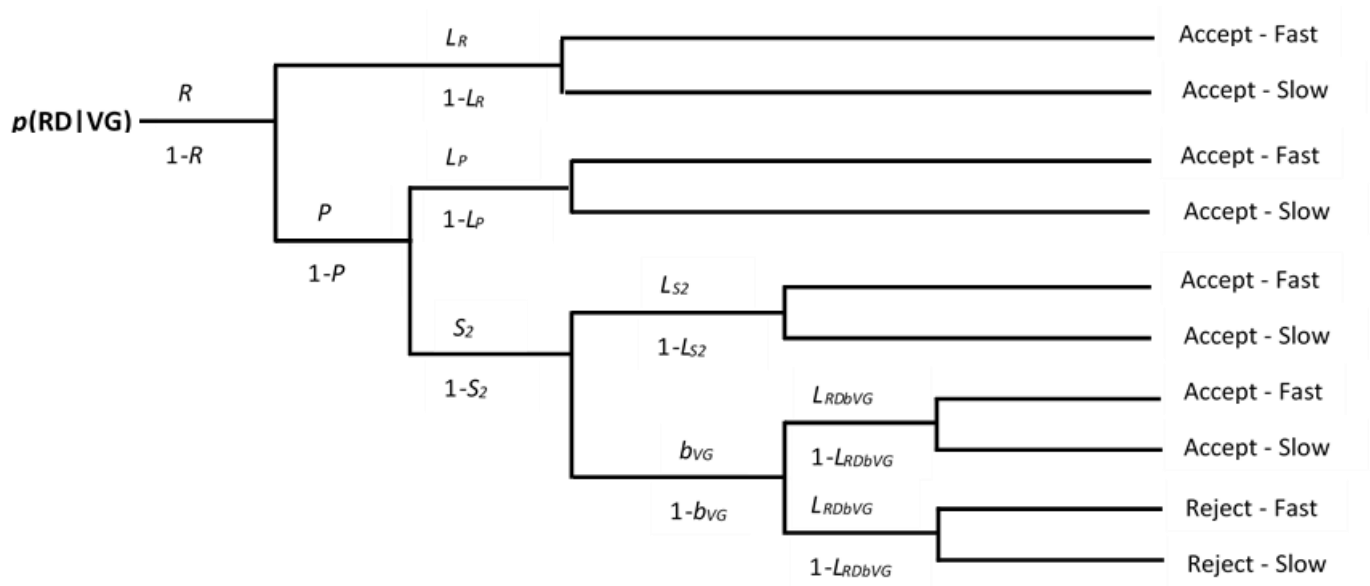


Figure 3.2. Latency extended conjoint recognition model trees for related distractors.

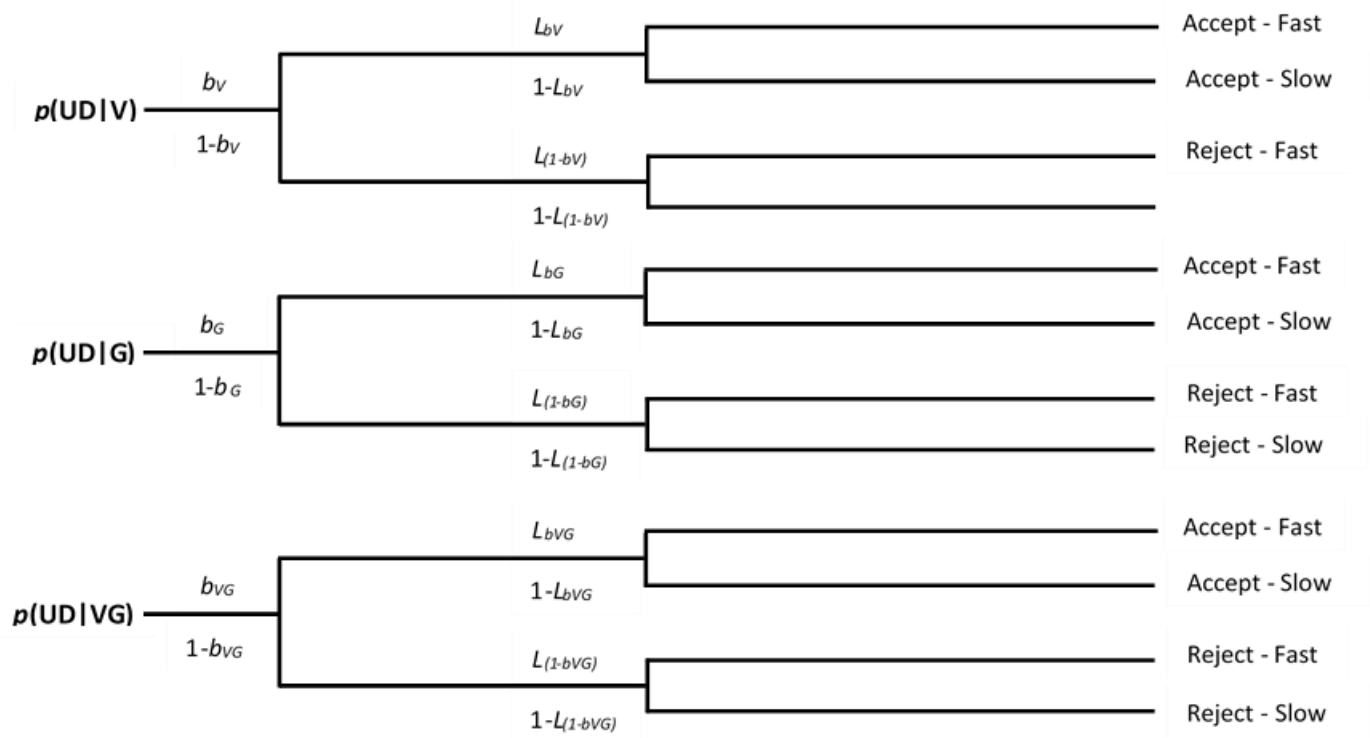


Figure 3.3. Latency extended conjoint recognition model trees for unrelated distractors.

Dynamic Model of Memory

I addressed how the new latency extended conjoint recognition model could directly measure memory process speeds and thus answer theoretical questions on their relative order, but it is important to discuss how this new model compares with some contemporary memory models that also incorporate the time-course of retrieval processes. Cox and Shiffrin (2017) introduced the dynamic model of memory, which integrates information on how the changes in the familiarity index over time affects recognition judgments.

The dynamic model may be described as an extension of univariate models of memory. In this model, memory events contain two types of traces: content memory, which includes information specific to the event such as perceptual and semantic information, and context memory, which includes information surrounding the event such as its time, location, and temporal and positional relations to other events. Both content and context memory make up a single metric of memory strength, called "familiarity." Note that this is different than familiarity defined in conjoint recognition, where familiarity is strictly a gist process that isn't combined with target and context recollection, unlike how it's defined in the dynamic model.

Unlike classic univariate theories where positive recognition decisions are made on the basis of this familiarity reaching a certain threshold, the dynamic model emphasizes the changes of familiarity over time. When a test probe is presented at t_0, it initially only contains context features (i.e. the room the probe is presented in, where it is located on the screen, the time it's presented). The familiarity metric at t_0 therefore reflects the overall match between the test context and the study context of targets stored in memory. As time passes, content features stored in memory are sampled and matched with content features of the probe, and its familiarity increases or decreases depending on the results of the matching. For example, the average

familiarity of targets will increase over time because more and more content features match the probe, whereas the average familiarity of distractors will decrease over time because the sampled content will not match the probe.

There are two predictions the dynamic model makes that are relevant to the present study on process latencies. First, the dynamic model describes particular types of memory traces being available for memory decisions at different times, and that they are not dependent on manipulations that increase or decrease their strength (like target repetition). According to Cox and Shiffrin, semantic content is accessed earlier than information surrounding the modality (auditory or visual presentation) or the plurality of the target. Second, the inverted U function of false memory occurs because semantic content is accessed earlier than other traces. Familiarity to related distractors initially increases due to fast semantic content sampling, and this contributes to the increase in false acceptances in early deadlines. False acceptances decrease as deadlines decrease, because average familiarity decreases over time as more and more memory features mismatch the probe.

The prediction that semantic content is accessed quickly while other types of traces are accessed slowly is compatible to dual process predictions. As stated before, dual recollection theory describes two types of recollection, target and context recollection, and that these two recollections may have different speeds. Regarding the first point, dual recollection does not make any assumptions on whether or not certain memory processes are faster than others based on their strength. Regarding the second point, fast phantom recollection may be responsible for the initial increase in false alarms at shorter deadlines, and prior research has shown that phantom recollection increases with increased gist. It is possible, then, that it is not the case that semantic content is accessed quickly, but that the process that is affected by semantic content is

accessed quickly.

While both the dynamic model and the latency extended conjoint recognition model both provide a model-based explanation to the inverted U false memory curve, the latency extension model has two main advantages. First, the dynamic model still cannot directly measure process speeds and relies on inferences. The strength of the latency extension model is that it parameterizes process latencies, and allows statistical tests to be conducted to compare how fast one process is to another. Second, because the dynamic model combines all types of memory traces into a single familiarity metric, it is not possible to identify what types of traces are accessed earlier than others using the model alone. Latency extension models, on the other hand, are able to define the relative order of memory processes, and are therefore able to predict the kinds of manipulations that will be relevant in a recognition decision earlier or later than others. The fact that latency values can be measured directly in latency extended models is what makes it appealing.

CHAPTER 4

EXPERIMENT 2: TESTING THE LATENCY EXTENSION CONJOINT RECOGNITION MODEL

There were two goals for Experiment 2. The first goal was to apply the new latency modeling techniques to real data, and see if the extended model provides adequate fit to them. The second goal was to see if familiarity was faster than both forms of recollection: target and context. According to the dual recollection theory, it is possible that faster phantom (context) recollection is responsible for greater false memory during shorter response intervals, rather than familiarity. This is the first study to directly test the relative speeds of familiarity and the two forms of recollection against each other.

Experiment 3 is a study designed to specifically address some additional questions that arose from Experiment 2. There were two goals for Experiment 3. First, there may be some concern that imposing response deadlines affects the speed in which processes become available due to strategy. In order to determine if the results of Experiment 2 can be applied to false memory studies with response deadlines imposed on them, it is necessary to run the same latency extension analysis on a response deadline study. Second, analyzing response latency data in conjunction with the results from latency extension modeling may provide some insight into how fast the processes come online, rather than just their relative speeds.

Method

Participants

Five datasets included conjoint recognition data from 188 introductory psychology students. There were 19 participants in dataset 1, 28 participants in dataset 2, 59 participants in dataset 3, 31 participants in dataset 4, and 23 participants in dataset 5. One participant was

removed from dataset 5 for a total of 22 participants, because they did not follow the instructions for the task correctly. The students participated to fulfill a course requirement.

Materials

All five conditions used a standard conjoint recognition design using DRM lists (Deese, 1959; Roediger & McDermott, 1995). Lists were chosen from a pool of 200 DRM lists aggregated from three different studies (Arndt, 2012; Brainerd & Wright, 2005; Flegal Atkins Reuter-Lorenz, 2010). 75 lists that produce the highest levels of false memory were pulled from Stadler et al.'s (1999) lists for study as well as additional lists from Roediger et al. (2001). Each of the 75 lists were four words long. In dataset 1, 2, and 3, the four words were chosen randomly from the available DRM lists, and were not optimized on backward association strength (BAS). The average BAS for the study words in these experiments was 0.41. In order to elevate false memory levels, dataset 4 and dataset 5 used study words with the highest BAS, and the average BAS for these conditions was 0.46. Because the experiment was divided into three different study-test cycles, there were a total of 25 lists of four items per study. The lists for each study were distributed randomly.

Each test was comprised of 75 items of three different cue types: 25 targets that had been presented during study, 25 critical distractors that were related to targets but were never presented, and 25 unrelated distractors that were neither presented at study nor related to targets. Therefore, across all three tests, 255 items were tested. With the exception of dataset 4 and dataset 5, unrelated distractors were chosen from critical distractors from unused DRM lists. For dataset 4 and dataset 5, unrelated distractors were high frequency abstract words that were chosen from the Toglia & Battig (1978) word norms. The words that were chosen were not associated to any of the items presented at study.

Each cue word was factorially paired with one of three possible probes: “The word was presented on the list,” “The word was NOT presented but related to a word on the list,” “The word was NEITHER presented NOR related to a word on the list.” The conjoint recognition paradigm refers to these instructional conditions as “V: Verbatim”, “G: Gist”, and “VG: Verbatim or Gist” conditions. Participants were asked to accept or reject each probe depending on whether they judged the statement to be true or false. It is important to note that each study-test cycle was independent of each other, so that an item on the third test was never associated with a word on a previous test. Three versions of the test were created, with different randomizations of the study and test words in consideration of list-order effects.

Procedure

Participants for all conditions were informed that they would be taking part in a memory test and should try to remember the words presented to the best of their ability. They were given instructions on both the study and test portions of the experiment prior to the start of the study, and they were also given sample problems to make sure they understood the task. During the instructions, participants were warned that the memory test would include words that they had seen before as well as words that were new, and that they would have to make judgments on whether the episodic description attached to the cue was accurate. They were specifically instructed to only accept a probe if they believed them to be true, and to reject them otherwise. The experiments took place electronically on a computer screen.

For all conditions except dataset 3, words during study were presented at a 2-s rate. For dataset 3, they were presented at a 1-s rate. Words were presented centered on the screen and printed in 72-point font. Between each word, there was a 3-s inter stimulus interval where a black fixation cross was displayed on the center of the screen. Each study took approximately

8.3 minutes to complete, with the exception of dataset 3 which took 4.15 minutes to complete. After study, there was a brief 3-minute break where the participant was briefed again on the instructions of the upcoming test. The test portion followed if the participant had no questions or concerns. Test probes were presented one at a time, and each probe was presented for 4 seconds during which time the participant had to make a response with a key press of 'yes' or 'no.' Each test portion took approximately 30 minutes to complete. A summary of all of the differences and similarities across the five available conditions are provided on Table 4.1.

Table 4.1

Experiment 2 methodological differences

Dataset	*n*	Study	Average BAS	Unrelated Distractors
1	19	2s/word presentation	0.41	critical words from unused DRM lists
2	28	2s/word presentation	0.41	critical words from unused DRM lists
3	59	1s/word presentation	0.41	critical words from unused DRM lists
4	31	2s/word presentation	0.46	high frequency abstract words
5	23	2s/word presentation	0.46	high frequency abstract words

Latency Binning

Latency extension models require data to be binned by response latencies. That is, for every observable response probability that is included in an unextended model, they must be further discriminated by response latency. For the purposes of this study, two latency bins, 'fast' and 'slow', were created for each cue-probe combination and response (e.g. fast target accept in

the V condition, slow target accept in the V condition, fast target reject in the V condition…). Latency boundaries for binning responses were determined by taking the reaction times of all responses, log-transforming them, identifying the median, and then converting the median back to the raw latency value and using it to separate fast and slow bins. The binning methods were adopted from those discussed in Heck and Erdfelder (2016).

Results and Discussion

Descriptive Statistics

Descriptive stats for acceptance rates of each probe and descriptive stats for latencies of each probe is reported in Table 4.2. Although not the main interest of the current experiment, this data provides a check to make sure the subjects were understanding the task and that the data does not deviate from previous conjoint recognition data. In terms of the acceptance probabilities, the grand means were $p(\mathrm{VG}) = 0.78$, $p(\mathrm{V}) = 0.74$, and $p(\mathrm{G}) = 0.33$ for targets, $p(\mathrm{VG}) = .72$, $p(\mathrm{G}) = .61$, and $p(\mathrm{V}) = .42$ for related distractors, and $p(\mathrm{VG}) = .35$, $p(\mathrm{G}) = .31$, and $p(\mathrm{V}) = .20$ for unrelated distractors. To summarize, $p(\mathrm{VG}) > p(\mathrm{V}) > p(\mathrm{G})$ for targets, $p(\mathrm{VG}) > p(\mathrm{G}) > p(\mathrm{V})$ for related distractors, and $p(\mathrm{VG}) = p(\mathrm{G}) > p(\mathrm{V})$ for unrelated distractors (the difference between VG and G acceptance probabilities were not significantly different for unrelated distractors: mean difference=0.04, SE=0.02, p=0.10).

Table 4.2

Mean Acceptance Probabilities (SDs) and Mean Accept and Reject Latencies (SDs) for the 9 Item X Probe Combinations of Experiments 1-6

Item | probe type	Experiment					
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Mean
Probability:						
RD | V	.31(.46)	.34(.47)	.38(.49)	.51(.50)	.57(.49)	.42
RD | G	.53(.50)	.58(.49)	.63(.48)	.64(.48)	.65(.48)	.61
RD | VG	.70(.46)	.70(.46)	.74(.44)	.68(.47)	.77(.42)	.72
T | V	.68(.47)	.77(.42)	.78(.42)	.65(.48)	.80(.40)	.74
T | G	.29(.45)	.29(.46)	.30(.46)	.52(.50)	.25(.43)	.33
T | VG	.77(.42)	.78(.42)	.82(.39)	.68(.47)	.83(.37)	.78
UD | V	.16(.37)	.19(.39)	.16(.36)	.27(.44)	.20(.40)	.20
UD | G	.32(.47)	.29(.45)	.26(.44)	.34(.47)	.34(.47)	.31
UD | VG	.36(.48)	.34(.47)	.30(.46)	.37(.48)	.38(.49)	.35
Latency-accept:						
RD | V	1.80(.69)	1.82(.66)	2.10(.77)	2.20(.71)	2.08(.83)	2.00
RD | G	2.06(.66)	2.15(.59)	2.31(.70)	2.25(.75)	2.36(.75)	2.23
RD | VG	1.86(.60)	1.92(.61)	2.06(.67)	2.24(.74)	2.04(.71)	2.02
T | V	1.51(.56)	1.59(.54)	1.69(.68)	1.90(.77)	1.71(.69)	1.68
T | G	1.89(.75)	1.99(.65)	2.20(.83)	2.12(.80)	2.23(.73)	2.01
T | VG	1.67(.54)	1.72(.57)	1.81(.66)	2.02(.78)	1.75(.62)	1.79
UD | V	1.65(.64)	1.83(.60)	2.10(.77)	2.35(.76)	2.10(.82)	2.01
UD | G	2.04(.66)	2.22(.57)	2.41(.72)	2.46(.77)	2.33(.77)	2.29
UD | VG	1.88(.65)	1.97(.59)	2.22(.73)	2.29(.78)	2.24(.74)	2.12
Latency-reject:						
RD | V	1.81(.60)	1.94(.57)	2.17(.68)	2.28(.73)	2.13(.69)	2.07
RD | G	2.03(.76)	2.05(.66)	2.35(.73)	2.39(.80)	2.41(.74)	2.25
RD | VG	1.72(.73)	1.91(.67)	2.19(.74)	2.33(.74)	2.29(.76)	2.09
T | V	1.67(.75)	1.85(.59)	2.08(.69)	2.24(.69)	2.12(.71)	1.99
T | G	1.94(.67)	2.01(.57)	2.23(.71)	2.41(.73)	2.20(.72)	2.16
T | VG	1.69(.79)	1.89(.59)	2.25(.69)	2.28(.77)	2.10(.73)	2.04
UD | V	1.70(.59)	1.81(.55)	1.93(.65)	2.21(.72)	1.90(.64)	1.91
UD | G	2.02(.71)	2.08(.62)	2.35(.73)	2.41(.73)	2.39(.75)	2.25
UD | VG	1.99(.72)	2.01(.67)	2.21(.71)	2.33(.74)	2.12(.72)	2.13

The distribution of responses parallel what has been found in prior conjoint recognition experiments. Brainerd et al. (2014) reviewed a corpus of 297 conjoint recognition data sets using semantically related and unrelated word lists like the methods in this experiment. They found p(VG) > p(V) > p(G) for targets, p(VG) > p(G) > p(V) for related distractors, and p(VG) = p(G) > p(V) for unrelated distractors. Therefore, the conjoint recognition data in the present experiment behaved in a way expected out of these kinds of experiments.

Core Model Fit

Before implementing the latency extension to the conjoint recognition model, it is necessary to first see if the model provides acceptable fits to the data. The conjoint recognition model contains three parameters for targets (I, E, S_T), three parameters for distractors (R, P, S_D), and three parameters for bias (b_v, b_g, b_{vg}) for a total of nine. The present design provides nine free probabilities with which to estimate them. Because erroneous recollection rejection generally has very low values, E was assumed to be zero in order to have an unsaturated model with which to run fit tests. Therefore, the fit test for this model to determine if it is acceptable to account for the response probabilities is a $G^2(1)$ statistic, with a critical value 3.84 to reject the null hypothesis of fit at the 0.05 level. All five datasets provided acceptable fits to this model, and the null hypothesis could not be rejected. Fit statistics and parameter estimates of the core model can be found on (Table 4.3).

Table 4.3

Maximum likelihood estimates of the Conjoint Recognition Model's retrieval and bias parameters for Experiment 2 datasets

Dataset			Retrieval-targets		Retrieval-related distractors:			Bias		
	G^2	df	I	S_1	R	P	S_2	b_V	b_G	b_{VG}
1	.24	1	.46	.31	.28	.21	.15	.16	.32	.37
2	2.60	1	.49	.40	.29	.14	.26	.19	.29	.33
3	.02	1	.51	.46	.32	.12	.39	.16	.26	.30
4	.44	1	.16	.42	.12	.03	.41	.27	.34	.36
5	.08	1	.60	.40	.14	.14	.56	.18	.33	.38

Latency Extension Model Fit

The core model is a 3 (probe: V, G, VG) × 3 (cue: target, related distractor, unrelated distractor) factorial structure with a total of 9 different conditions. In each of these 9 conditions, responses are binned into four different probability outcomes: accept fast, accept slow, reject fast, reject slow. Of these four probabilities, three are free to vary, and thus the latency extension model has 27 free empirical probabilities. Five degrees of freedom are used for latency parameters associated with each of the retrieval parameters, (L_I, L_{S1}, L_R, L_P, L_{S2}). Twelve more degrees of freedom are used for the latency parameters associated with various bias parameters, for a total of 17 degrees of freedom. With these constraints, the test of fit is a $G^2(10)$ statistic that is asymptotically distributed as X^2 with a critical value of 18.31 to reject the null hypothesis at the .05 level.

Results of the latency extension are summarized on Table 4.4.

Table 4.4
Maximum likelihood estimates of the Conjoint Recognition Model's latency parameters for Experiment 2 datasets

Dataset			Latency-retrieval					
	G^2	df	L_I	L_{S1}	L_R	L_P	L_{S2}	
1	7.79	10	.72	.69	.30	.54	.47	
2	11.44	10	.71	.67	.44	.86	.42	
3	62.42	10	.72	.67	.50	1.00	.37	
4	2.49	10	.71	.63	.47	1.00	.43	
5	36.71	10	.71	.70	.52	.89	.53	
Mean			.71	.67	.45	.86	.44	
			Latency-list items			Latency-related distractors		
			L_{TbV}	L_{TbG}	L_{TbVG}	L_{RDbV}	L_{RDbG}	L_{RDbVG}
1	7.79	10	.63	.02	.69	.68	.50	.64
2	11.44	10	.57	.00	.54	.50	.33	.52
3	62.42	10	.53	.00	.43	.40	.15	.44
4	2.49	10	.54	.00	.53	.45	.00	.44
5	36.71	10	.54	.32	.49	.44	.40	.42
Mean			.56	.07	.54	.49	.28	.49
			Latency-bias (accept unrelated distractors)			Latency-bias (reject unrelated distractors)		
			L_{bV}	L_{bG}	L_{bVG}	$L_{(1-bV)}$	$L_{(1-bG)}$	$L_{(1-bVG)}$
1	7.79	10	.60	.31	.40	.63	.44	.43
2	11.44	10	.53	.28	.44	.58	.38	.42
3	62.42	10	.50	.30	.44	.62	.34	.44
4	2.49	10	.67	.28	.38	.59	.33	.44
5	36.71	10	.42	.35	.49	.49	.39	.43
Mean			.54	.30	.43	.58	.38	.43

Model fits were adequate for three out of five data sets. Similar to procedures in Experiment 1, a test for individual differences was conducted for the condition where fit failed. The test for individual differences is a X^2 statistic of the following form (Smith & Batchelder, 2008):

$$X^2(N-1) = \sum_{i=1}^{N}\{(R_i - R_e)^2 / R_e\} + (M - R_i - R_e^*)^2 / R_e^*\} \qquad (16)$$

Each individual differences test is applied to each cue-probe combination at a particular latency speed, for a total of 36 tests per data set. In the above equation, N is equal to the number of subjects. *M* is the total number of examples of a particular cue-probe combination that

participants respond to (e.g. target in the V condition). R_i is the number of cases a particular cue-probe combination at a particular latency with a particular response was observed for a participant (e.g. fast accept of a target in the V condition). R_e is $\sum_i^N [R_i]$/ N, and $R_e* = M - R_e$. For N = 25, the critical value of X^2 to reject the null hypothesis of no individual differences at the .05 level was 37.70.

For both dataset 3 and dataset 5, significant individual differences were observed in the data, with dataset 3 failing all 36 tests and dataset 5 failing all but two. Therefore, because it is likely that model fits failed for these two experiments because of large individual differences between subjects, it is necessary to move on to hierarchical models which may account for them (Pratte & Rouder, 2010). Hierarchical model analysis involves running the latency-extended conjoint recognition model for each individual subject and observing fits. The average fit for all 59 subjects in dataset 3 for 10 degrees of freedom was 13.33, with 79.66% of the subjects showing adequate fit. The average fit for 22 subjects in dataset 5 for 10 degrees of freedom was 11.56, with 86.36% of the subjects showing adequate fit. Therefore, we can conclude that under the hierarchical model, fit was adequate for both dataset 3 and dataset 5, and that fit failed not because the model was inaccurate but because there were large individual differences in the data. Further discussion regarding the parameter estimates was therefore made under the assumption that the initial latency extension model was correct.

Process Speeds

Likelihood ratio tests (LRT) were run to determine if parameter estimates were significantly different from each other in each dataset. In these tests, we constrain the model being used so that the two parameters we want to test against each other are equal. We then take a look at the G^2 value and compare model fits of this new constrained model to the unconstrained

model. The test statistic is computed by subtracting the fit statistic of the unconstrained model from the constrained model, and the resulting difference is a statistic asymptotically distributed as X^2, with *df* equal to the difference in free parameters between the two models (for the current experiments, *df* = 1).

There were three notable results from the significance tests to determine the relative speed of memory processes compared to each other. For retrieval processes involving targets, there were no significant differences in process speeds for identity judgement, *I*, and familiarity, S_1, for all experiments. For retrieval processes involving related distractors, phantom recollection, *P*, was faster than recollection rejection, *R*, for all of the datasets except for dataset 4 where there was no significant difference (dataset 1: $G^2(1)=8.08$, dataset 2: $G^2(1) = 11.19$, dataset 3: $G^2(1)=31.00$, dataset 5: $G^2(1)=21.05$). *P* was also faster than familiarity, S_2, in two of the datasets (dataset 3: $G^2(1)=24.14$, dataset 5: $G^2(1)=15.98$). Finally, there was no significant difference in the speed of *R* and S_2 in all of the datasets except for dataset 3, where it was faster $G^2(1)=5.39$. In summary, identity judgment and familiarity for targets become available at around the same speed, and phantom recollection is the fastest process for related distractors compared to familiarity and recollection rejection.

The same type of analysis was conducted on latency for bias for accepting and rejecting unpresented unrelated distractors to see if there were bias-related differences in accepting or rejecting items across conditions (V, G, VG) when memory process-related traces cannot be retrieved. Latency of the bias parameter on accept conditions were faster in the V condition than the G condition in all of the datasets (dataset 1: $G^2(1)=16.43$, dataset 2: $G^2(1) = 17.39$, dataset 3: $G^2(1)=23.28$, dataset 5: $G^2(1)=34.71$) except for dataset 4. They were also faster in the V condition than the VG condition for dataset 1 ($G^2(1)=8.02$) and dataset 5 ($G^2(1)=19.63$), but were

not significantly different for the others. Latency of the bias parameter on accept conditions were faster in the VG condition compared to the G condition for all datasets (dataset 2: $G^2(1)=10.69$, dataset 3: $G^2(1)=16.88$, dataset 4: $G^2(1)=10.09$, dataset 5: $G^2(1)=4.20$) except for dataset 1 where there was no significant difference. Therefore, the general pattern is that responses for accepting an unpresented distractor on the basis of bias was fastest for the V condition and slowest for the G condition.

Latency of the bias parameters on reject conditions also painted a similar picture. Latency of the bias parameter on reject conditions were faster in the V condition than the G condition on all of the datasets (dataset 1: $G^2(1)=24.71$, dataset 2: $G^2(1)=35.26$, dataset 3: $G^2(1)=160.05$, dataset 4: $G^2(1)=8.93$, dataset 5: $G^2(1)=51.56$). Rejecting bias was also faster in the V condition compared to the VG condition in all datasets except for dataset 4 (dataset 1: $G^2(1)=27.11$, dataset 2: $G^2(1)=22.02$, dataset 3: $G^2(1)=66.34$, dataset 5: $G^2(1)=16.59$). Last, rejecting bias in the VG condition was faster than in the G condition for dataset 3 ($G^2(1)=18.36$) and dataset 5 ($G^2(1)=8.87$), and no significant difference in the other datasets. Again, V conditions had the fastest latency times and G conditions had the slowest.

To summarize, consistent patterns emerged across all 6 conjoint recognition experiments for process speeds. For retrieval processes for targets, identity judgment was roughly the same speed as familiarity. For related distractors, phantom recollection was the fastest for all experiments, while recollection rejection and familiarity were the slowest. For bias parameters, rejecting an item was generally faster than accepting an item, and responses to items in the V condition and VG condition were fastest while responses to items in the G condition were the slowest. The finding that phantom recollection is faster than familiarity supports the dual recollection hypothesis that faster phantom (context) recollection is responsible for the left arm

of the inverted U function of false alarm over time. This is different than prior assumptions in the literature that familiarity is always faster than recollective processes.

CHAPTER 5

EXPERIMENT 3: PHANTOM RECOLLECTION AS A FAST PROCESS

We saw in Experiment 2 that a latency extension can be applied to the conjoint recognition model with acceptable fits, and that it provided evidence that a fast recollective process (phantom recollection) and not familiarity was responsible for the inverted U curve of false memory. The goal of Experiment 3 was to test two follow-up questions from Experiment 2. First, are the relative speed of memory processes consistent when we impose response deadlines in recognition experiments, or are they dependent on this experimental manipulation? Second, if phantom recollection is the fast process responsible for elevated false memory at fast response latencies, then can we observe greater values of phantom recollection compared to familiarity at earlier deadlines?

Method

Participants

189 introductory psychology students were recruited for course credit. Each participant was randomly selected to one of three conjoint recognition conditions (V, G, and VG), and for two response time conditions (timed and untimed). For the timed conditions, there were 37 participants in the V condition, 36 in the G condition, and 33 in the VG condition. For the untimed conditions, there were 27 participants in the V condition, 28 in the G condition, and 28 in the VG condition.

Materials

The design was similar to the format of Experiment 2 with a few key changes. First, the conjoint recognition instructional condition was between-subjects (for V, G, VG). The conjoint recognition instructions were the same as those presented in Experiment 2. Second, whether

subjects answered test probes under a response deadline or not was also a between-subject manipulation. 30 DRM lists were selected from the Roediger et al. (2001) norms, and the lists that produced the highest false memories were identified and prioritized using Stadler's (1999) norms. From these 30 lists, 12 words with the highest BAS were chosen as targets for the study list, and the mean BAS was 0.34.

For the test list, there were 30 critical distractors, one each from the 30 DRM lists that were presented during study. 30 targets were chosen randomly, one from each list. Last, 30 unrelated distractors were chosen from critical distractors from unused lists from the Roediger norms. Because there weren't enough DRM lists in the 12-word DRM norms, the remaining words were chosen from a large pool of 200 DRM lists aggregated from three different studies that were used in Experiment 2 (Flegal Atkins Reuter-Lorenz, 2010; Arndt, 2012; Brainerd & Wright, 2005).

Like Experiment 2, Experiment 3 also had three different study-test cycles, and like Experiment 2, each study-test cycle was independent of each other. Therefore, there were 10 DRM lists (120 words) per study, and 10 targets, 10 critical distractors, and 10 unrelated distractors per test. Three different test versions were created with different DRM lists selected randomly to appear on either the first, second, or third study-test cycle. For the timed condition, there were five different response deadlines for the test: 100ms, 300ms, 500ms, 750ms, 1000ms. Therefore, there were 6 test items for each deadline condition (e.g. 6 targets at 100ms, 6 targets at 300ms… 6 unrelated distractors at 1000ms).

Procedure

At the beginning of the experiment, participants were told that they would be taking part in a memory test and should try to remember the words presented to the best of

their ability. For the study portion, words were presented centered on the screen and printed in 72-point font at a 2-s rate. Between each study word a fixation cross was presented centered on the screen for 200ms. Each study took approximately 4 minutes to complete.

After study, they were given detailed instructions of the memory test, including examples of the conjoint recognition instructions to make sure they understood the task. They were warned that the memory test would include words that they had seen, as well as words that were new, and that they would have to make judgments on whether or not the conjoint recognition instructional condition they were assigned to (V, G, or VG) reflected the cue accurately. They were specifically instructed to only accept a probe if they believed them to be true, and to reject them otherwise. Participants were also given a short five-word practice test so that they would know what to expect regarding response times.

In the timed condition, participants were instructed that during the test, a word would be presented for a brief period, followed by a tone indicating that they could make their memory response. It was stressed that they should only make responses after the tone, and that they should also respond as fast as possible immediately upon hearing it. Responses were made by either pressing a 'yes' or 'no' key on a keyboard. Depending on the condition, the tone would be played at 100ms, 300ms, 500ms, 750ms, or 1000ms after the cue was presented on the screen, after which the word would disappear. If a response key was pressed within 300ms after the tone, participants saw a green colored text with their reaction time. If they responded after 300ms, they saw a red colored text with a message telling them that they were too slow.

The untimed condition was the same as the timed condition, except with two changes reflecting the lack of deadlines. First, there was no tone indicating that participants should respond, and they were not instructed about this tone. Second, they received no feedback on

how quickly they made a response, and instead saw a black fixation cross centered on the screen between each test item. The test was completed in a self-paced manner similar to standard DRM experiments without response deadlines.

Latency Binning

The same latency binning procedure used in Experiment 2 was used for Experiment 3. Because there were six between-subject response deadline conditions (100ms, 300ms, 500ms, 750ms, 1000ms, untimed), six different latency boundaries were calculated for each condition separately. For each deadline condition, reaction times of all responses were log-transformed, the median of the log-transformed reaction times was identified, and then the median was converted back to the raw latency value and was used as the boundary to separate fast and slow bins. Latency boundaries were calculated separately for each condition to avoid ceiling and floor effects for responses in the extreme deadline conditions (e.g. 100ms and 1000ms), where all responses are binned as fast or slow, and thus no estimates can be made for process latencies.

Results and Discussion

Descriptive Statistics

Descriptive statistics for acceptance rates of each probe and descriptive statistics for latencies of each probe are reported in Table 5.1. Similar to Experiment 2, acceptance probabilities were well-behaved, indicating that the subjects understood the task properly. The grand means for the six timing conditions were p(VG) = .77, p(V) = .60, and p(G) = .47 for targets, p(VG) = .81, p(G) = .65, and p(V) = .40 for related distractors, and p(VG) = .26, p(G) = .30, and p(V) = .11 for unrelated distractors. To summarize, p(VG) > p(V) > p(G) for targets, p(VG) > p(G) > p(V) for related distractors, and p(VG) = p(G) > p(V) for unrelated distractors (the difference between VG and G acceptance probabilities were not significantly different for

unrelated distractors: mean difference=.04, SE=.05, p=.41). Again, the distributions in Experiment 3 paralleled prior distributions of acceptance probabilities in conjoint recognition experiments.

Bias corrected (two-high-threshold method; 2HT) acceptance rates for targets and related distractors are reported on Table 5.2. To see if there were any trends in acceptance rates across the five timed conditions, a 3 (instruction: V/G/VG) × 5 (response deadline: 100ms/300ms/500ms/750ms/1000ms) × 2 (probe: targets vs. related distractors) ANOVA was conducted on bias corrected acceptance rates for targets and related distractors. There was a main effect of Deadline, $F(4, 412) = 7.88$, $p<0.001$, partial $\eta^2 = 0.07$, where longer deadlines had higher acceptance rates than shorter deadlines. There was also a main effect of Instruction, $F(1, 103) = 19.76$, $p<0.001$, partial $\eta^2 = 0.28$, where VG instructions had higher acceptance rates than V or G instructions, and V instructions had higher acceptance rates than G instructions.

Three significant interactions were observed. There was a two-way Probe × Instruction interaction, $F(2, 103)=39.26$, $p<0.001$, partial $\eta^2 = 0.43$. For targets, there was a greater acceptance rate for V and VG instructional conditions compared to the G condition. For related distractors, however, there was a greater acceptance rate for the VG condition compared to both V and G conditions, but there was no difference in acceptance rates between V and G conditions. There was also a two-way Deadline × Probe interaction, $F(4, 412)=6.84$, $p<0.001$, partial $\eta^2 =0.06$. While there was no significant difference in accepting targets across all timing conditions, accepting related distractors increased over time. Last, there was a three-way Deadline × Probe × Instruction interaction, $F(8, 412)=6.50$, $p<0.001$, partial $\eta^2 =0.11$. For V condition for targets, there was a greater acceptance rate with longer deadlines, but no such trend was observed for related distractors. For G and VG conditions the reverse was true, with greater

acceptance rate of related distractors with longer deadlines, but no such trend for targets.

Table 5.1

Mean Acceptance Probabilities (SDs) and Mean Accept and Reject Latencies (SDs) for the 9 Item X Probe Combinations of 6 deadline conditions

Item | probe type	Response Deadlines						
	100ms	300ms	500ms	750ms	1000ms	none	Mean
Probability:							
RD | V	.45(.28)	.46(.21)	.47(.27)	.39(.22)	.39(.26)	.23(.22)	.40
RD | G	.49(.21)	.65(.22)	.64(.18)	.76(.19)	.68(.23)	.70(.22)	.65
RD | VG	.77(.18)	.81(.18)	.80(.21)	.83(.14)	.81(.23)	.85(.18)	.81
T | V	.53(.26)	.58(.16)	.62(.25)	.52(.17)	.64(.21)	.68(.18)	.60
T | G	.52(.22)	.52(.20)	.54(.23)	.38(.20)	.38(.22)	.28(.18)	.47
T | VG	.78(.21)	.80(.14)	.78(.22)	.70(.18)	.77(.20)	.81(.20)	.77
UD | V	.16(.20)	.17(.21)	.13(.15)	.07(.14)	.08(.13)	.05(.10)	.11
UD | G	.35(.21)	.33(.24)	.41(.20)	.24(.22)	.28(.23)	.21(.20)	.30
UD | VG	.38(.30)	.32(.24)	.31(.25)	.16(.21)	.26(.19)	.13(.15)	.26
Latency-accept (in seconds):							
RD | V	.52(.27)	.61(.08)	.90(.38)	1.16(.79)	1.28(.18)	1.67(.77)	1.02
RD | G	.49(.11)	.61(.08)	.86(.22)	1.07(.26)	1.29(.15)	1.95(.54)	1.05
RD | VG	.45(.08)	.61(.15)	.79(.19)	1.01(.15)	1.33(.31)	1.32(.38)	.92
T | V	.50(.19)	.57(.09)	.79(.26)	1.03(.18)	1.40(.61)	1.51(.48)	.97
T | G	.47(.09)	.62(.15)	.82(.15)	1.11(.28)	1.28(.21)	2.32(1.00)	1.10
T | VG	.44(.08)	.62(.16)	.77(.09)	1.09(.20)	1.28(.16)	1.48(.52)	.95
UD | V	.49(.11)	.62(.20)	.78(.31)	1.61(1.44)	1.42(.47)	1.61(.79)	1.09
UD | G	.48(.11)	.60(.12)	.81(.17)	1.03(.14)	1.29(.16)	2.27(.85)	1.08
UD | VG	.45(.12)	.63(.18)	.98(.87)	1.10(.24)	1.34(.37)	1.70(.70)	1.03
Latency-reject (in seconds):							
RD | V	.50(.12)	.62(.09)	.80(.15)	1.02(.11)	1.31(.30)	1.68(.61)	.99
RD | G	.49(.11)	.62(.11)	.80(.16)	1.01(.33)	1.24(.13)	2.10(.88)	1.04
RD | VG	.47(.15)	.61(.13)	.84(.18)	1.70(2.31)	1.27(.12)	1.89(.87)	1.13
T | V	.48(.14)	.62(.15)	.81(.17)	1.26(.34)	1.26(.17)	1.78(.94)	1.04
T | G	.52(.12)	.64(.14)	.77(.10)	1.29(.58)	1.32(.32)	2.01(.56)	1.09
T | VG	.44(.09)	.62(.10)	.77(.11)	1.33(1.22)	1.29(.18)	2.00(.84)	1.08
UD | V	.48(.09)	.62(.20)	.77(.14)	1.04(.19)	1.29(.17)	1.50(.44)	.95
UD | G	.51(.11)	.60(.12)	.84(.25)	1.09(.48)	1.35(.22)	2.30(.70)	1.12
UD | VG	.51(.08)	.63(.18)	.77(.09)	1.01(.10)	1.30(.15)	1.93(.72)	1.03

Table 5.2

Bias-Corrected Mean Acceptance Probabilities (SDs) for the 9 Item X Probe Combinations of 6 deadline conditions

Item | probe type	Response Deadlines						
	100ms	300ms	500ms	750ms	1000ms	none	Mean
Probability:							
RD | V	.29(.32)	.29(.28)	.34(.29)	.32(.21)	.31(.23)	.18(.22)	.29
RD | G	.14(.24)	.31(.31)	.24(.29)	.52(.27)	.41(.32)	.49(.31)	.35
RD | VG	.39(.31)	.49(.30)	.49(.28)	.66(.25)	.56(.34)	.72(.27)	.55
T | V	.36(.31)	.41(.29)	.49(.33)	.45(.22)	.55(.26)	.63(.21)	.48
T | G	.17(.31)	.19(.22)	.13(.28)	.14(.29)	.10(.25)	.08(.26)	.14
T | VG	.40(.37)	.48(.28)	.47(.36)	.53(.28)	.52(.30)	.68(.28)	.51

Although results from the ANOVA do not show significant differences in false alarm acceptance rates in the three-way interaction, the shape of the trend is still consistent with prior literature in that false-positives to probes follows an inverted U-shape function. For related distractors, bias-corrected acceptance rates in the V condition increases, and then decreases after deadlines longer than 500ms (100ms=0.29, 300ms=0.29, 500ms=0.34, 750ms=0.32, 1000ms=0.31). It is also worth noting that incorrectly accepting targets in the G condition follows a similar inverted-U curve as well (100ms=0.17, 300ms=0.19, 500ms=0.13, 750ms=0.14, 1000ms=0.10). Key interest would be to see if changes in parameter estimates occur after the 500ms deadline condition.

Core Model Fit

The conjoint recognition model was fit to the six deadline conditions to see if the base model provided acceptable fits to the data. In Experiment 2, the model was made estimable by assuming that erroneous recollection rejection, *E*, was inconsequential and therefore set to zero. In the present experiment, however, this assumption could not be made, as the values for *E* were consistently high and thus resulted in poor fit. Out of the six deadline conditions, only one provided acceptable fits (1000ms condition; $G^2(1)=2.37$), and fit for an omnibus test with 6 degrees of freedom and critical value of 12.59 also failed ($G^2(6)=59.92$).

Another assumption that is often made with the conjoint recognition model is to assume that the bias parameters for the G and VG conditions are equal. Using this adjustment, fits were acceptable for all conditions except for the untimed condition. The omnibus test with 6 degrees of freedom and critical value of 12.59 failed when the untimed condition was included ($G^2(6)=19.33$), but was acceptable when only looking at conditions where participants were under a deadline, for a test with 5 degrees of freedom and a critical value of 11.07 ($G^2(5)=8.27$).

Running a test of individual differences to see if large variations in responses across individual subjects contributed to fit failure in the untimed condition revealed that subjects in the untimed condition had significant differences in responses for all instructional conditions. A hierarchical model analysis could not be implemented for Experiment 3 because conjoint recognition instructional conditions were run between subjects. Therefore, the bulk of the latency extension modeling results will focus on the conditions where response deadlines were imposed.

Latency Extension Model Fit

The core model is a 3 (probe: V, G, VG) × 3 (cue: target, related distractor, unrelated distractor) factorial structure with a total of 9 different conditions. In each of these 9 conditions, responses are binned into four different probability outcomes: accept fast, accept slow, reject fast, reject slow. Of these four probabilities, three are free to vary, and thus the latency extension model has 27 free empirical probabilities. Six degrees of freedom are used for latency parameters associated with each of the retrieval parameters, (L_I, L_E, L_{S1}, L_R, L_P, L_{S2}), and twelve are used for latency parameters associated with various bias parameters for a total of 9 degrees of freedom left for model fitting. Therefore, the test of fit is a $G^2(9)$ statistic that is asymptotically distributed as X^2 with a critical value of 16.92 to reject the null hypothesis at the .05 level. Results of the latency extension are summarized on Table 6.4. Model fits were adequate for all five deadline conditions.

Core Model Parameter Estimates

Parameter estimates for the core model are reported on Table 5.3. LRTs were run to determine if parameter estimates were significantly different from each other in each deadline, and for the current experiment, *df*=1 with a critical value of 3.84 to reject the null hypothesis.

Table 5.3
Maximum likelihood estimates of the Conjoint Recognition Model's retrieval and bias parameters for Experiment 3

Deadline			Retrieval-targets			Retrieval-related distractors:			Bias	
	G^2	df	*I*	*E*	S_1	*R*	*P*	S_2	b_V	$b_{G\text{-}VG}$
100ms	0.44	1	.26	.24	.38	.25	.37	.23	.16	.36
300ms	0.16	1	.27	.23	.46	.31	.23	.47	.17	.33
500ms	3.72	1	.24	.11	.49	.26	.21	.46	.13	.36
750ms	3.70	1	.31	.18	.32	.41	.11	.59	.07	.20
1000ms	0.25	1	.39	.13	.41	.38	.19	.49	.08	.27

For the three retrieval processes for targets, there were two deadlines in which S_1 was significantly greater than *I* (300ms: $G^2(1)=4.32$, 500ms: $G^2(1)=9.65$), two deadlines in which S_1 was significantly greater than *E* (300ms: $G^2(1)=3.86$, 500ms: $G^2(1)=11.70$), and three deadlines in which *I* was significantly greater than *E* (500ms: $G^2(1)=4.56$, 750ms: $G^2(1)=4.23$, 1000ms: $G^2(1)=12.39$). These patterns are consistent with prior experiments where models with *E*=0 was used, because while *E* could not be disregarded for the present study, it was still lower than the other memory processes involved in target recollection. It is also interesting to note that S_1 is greater than *I* in the two deadlines before 750ms. S_1 is a process that supports false alarms for targets, and incorrectly accepting targets for the G condition increased up to 500ms and then decreased after. It is possible that greater values of S_1 at these earlier deadlines contributed to the left arm of the inverted U function in falsely accepting targets in the G condition over time.

For the three retrieval processes for related distractors, *P* was greater than *R* for the fastest deadline (100ms: $G^2(1)=5.47$), but was smaller for the two longest deadlines (750ms: $G^2(1)=21.09$, 1000ms: $G^2(1)=8.93$). S_2 was greater than *P* for all deadlines except the fastest (300ms: $G^2(1)=4.14$, 500ms: $G^2(1)=5.26$, 750ms: $G^2(1)=14.87$, 1000ms: $G^2(1)=5.52$), and was

smaller than *R* for two deadlines (500ms: $G^2(1)=5.27$, 750ms: $G^2(1)=5.86$). The results confirm predictions from Experiment 2 that phantom recollection is greater than the other retrieval processes at fast deadlines, and was therefore responsible for the left arm of the inverted U function in false alarms over time. Familiarity also increases above phantom recollection by the 300ms deadline, further contributing to the left arm of the inverted U, before false alarms begin to decrease after the 500ms mark with recollection rejection increasing above familiarity and phantom recollection. This fits with the observation that false alarm rates appear to decrease after the 500ms deadline.

Levels of parameter estimates differed across the five deadlines. An omnibus test with a null hypothesis 100ms=300ms=500ms=750ms=1000ms with a $G^2(4)$ statistic with a critical value of 9.49 was run for each of the 6 memory parameters (I, E, S_1, P, R, S_2) and two bias parameters (b_V, $b_{G\text{-}VG}$). This test produced null hypothesis rejections for three parameters: S_2 (22.33), b_V (17.56), and $b_{G\text{-}VG}$ (38.41). Therefore, there is a significant trend in the increase of S_2 as response deadlines get longer, and a significant trend in the decrease of both bias parameters as response deadlines get longer. Although there were no significant differences in the levels of the other parameters, it is still worth noting that there were general linear trends as deadlines increased. For example for targets, *E* seemed to steadily decrease with longer deadlines, while *I* seemed to steadily increase with longer deadlines. For related distractors, *P* seemed to steadily decrease with longer deadlines, while *R* seemed to steadily increase with longer deadlines. There is a parallel decrease in *E* for targets and *P* for related distractors (Figure 5.1) that suggest that they both measure the same context recollection process. Visual representations of the trends are reported in Figure 5.1.

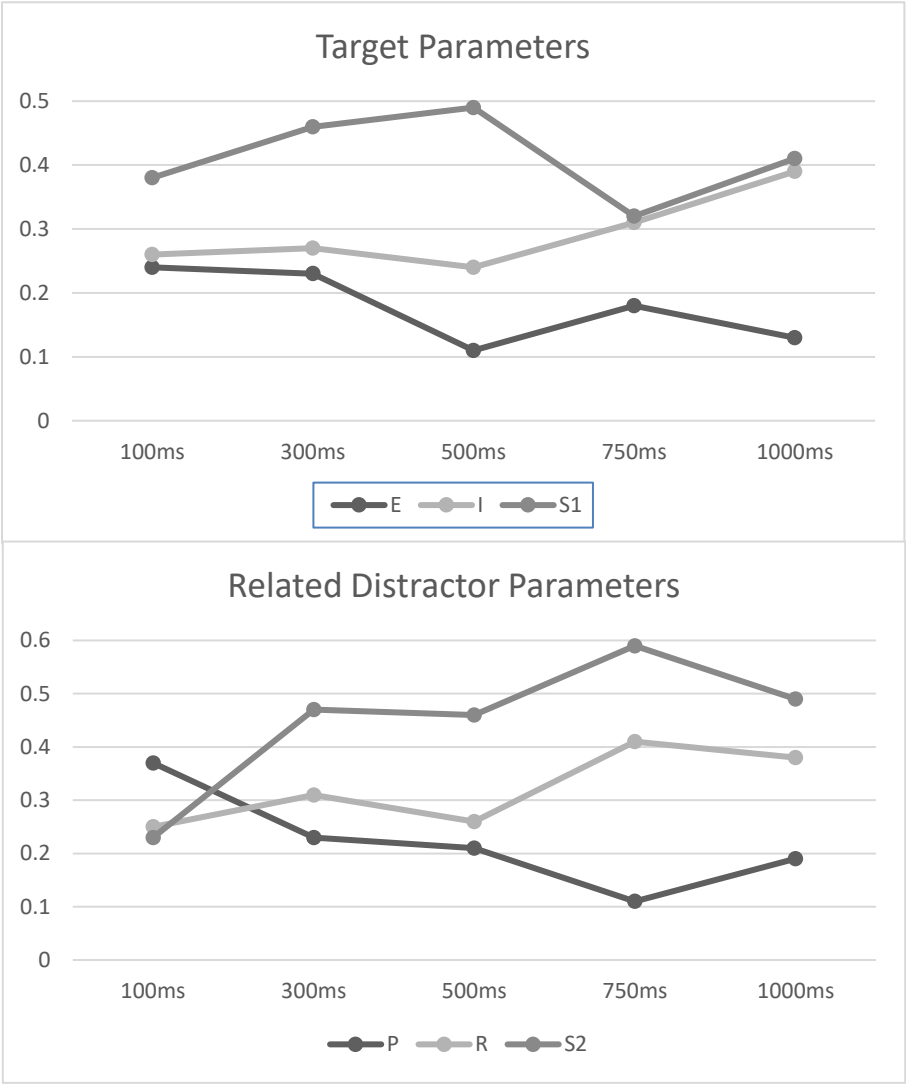


Figure 5.1. Parameter estimates for target and related distractor retrieval processes across deadline conditions.

Table 5.4

Maximum likelihood estimates of the Conjoint Recognition Model's latency parameters for Experiment 3

Dataset			Latency-retrieval					
	G^2	df	L_I	L_E	L_{S1}	L_R	L_P	L_{S2}
100ms	0.54	9	.68	1.00	.30	.85	.81	.10
300ms	0.24	9	.68	.84	.56	.72	.88	.27
500ms	3.75	9	.61	.30	.50	.68	.54	.42
750ms	3.82	9	.66	.38	.70	.81	1.00	.56
1000ms	0.42	9	.64	.00	.71	.53	.41	.65
				Latency-list items			Latency-related distractors	
			L_{TbV}	L_{TbG}	L_{TbVG}	L_{RDbV}	L_{RDbG}	L_{RDbVG}
100ms	0.54	9	.35	.04	.57	.21	.03	.49
300ms	0.24	9	.35	.00	.28	.16	.06	.35
500ms	3.75	9	.57	.40	.42	.41	.28	.43
750ms	3.82	9	.34	.44	.40	.21	.43	.53
1000ms	0.42	9	.64	.64	.42	.25	.59	.32
				Latency-bias (accept unrelated distractors)			Latency-bias (reject unrelated distractors)	
			L_{bV}	L_{bG}	L_{bVG}	$L_{(1-bV)}$	$L_{(1-bG)}$	$L_{(1-bVG)}$
100ms	0.54	9	.44	.49	.67	.47	.40	.25
300ms	0.24	9	.39	.44	.50	.56	.49	.43
500ms	3.75	9	.64	.40	.51	.59	.47	.48
750ms	3.82	9	.38	.69	.50	.71	.64	.70
1000ms	0.42	9	.50	.49	.38	.52	.48	.37

Process Speeds

Parameter estimates of the latency extended model are reported on Table 5.4. LRTs were run to see if the latency parameter estimates in each of the five deadline conditions were significantly different from each other. Recall that in Experiment 2 where no deadline was imposed during the recognition task, two main patterns emerged: 1.) no differences were

observed between the relative speeds of identity judgment, *I*, and familiarity, S_1 and 2.) phantom recollection, *P*, was faster than familiarity, S_2, and recollection rejection, *R*. A consistent pattern emerged for the present experiment. Regarding identity judgment and familiarity, there were no significant differences in their relative speeds for all conditions, similar to results from Experiment 2. Regarding parameters for related distractors, phantom recollection was faster than familiarity in two deadline conditions (100ms: $G^2(1)=3.92$, 300ms: $G^2(1)=4.22$), and was not significantly different from familiarity in the remaining conditions. Unlike Experiment 2, where *P* was also significantly faster than *R*, none of the timed conditions had significant differences in *P* and *R*. Recollection rejection was faster than familiarity in one out of six deadline conditions (300ms: $G^2(1)=4.10$). The results are consistent with Experiment 2 in that identity judgment is the generally the same speed as familiarity, and phantom recollection is generally faster than both recollection rejection and familiarity.

Unlike in Experiment 2, the relative speed of *E* could be measured in Experiment 3, because the core model accounted for it. There were no significant differences in process speed for *E* compared to *I* or S_1 except for one deadline condition. In the 100ms deadline condition, *E* was faster than *I* ($G^2(1)=4.44$) and S_1 ($G^2(1)=5.05$). This is consistent with the finding that context recollection processes are faster than familiarity or target recollection.

There may be a concern that the results of the latency parameters are affected by the binning methods, or that they do not reflect latency but rather reflect the strength of the core memory processes. Both of these concerns can be refuted by the data. First, regarding the concern that the latency parameters are affected by the binning methods; even though the binning boundaries were calculated separately for each timing condition, the overall results of the relative order of memory processes were the same. If the results were affected by binning methods, then

we may expect to see more variability in the results for process speeds. Second, regarding the concern that the latency parameters reflect core memory parameter strength and not speed; if this were true, then the results of the core model parameter estimates should match with the results of the latency parameter estimates. However, we clearly do not see this in the data. For targets for example, E is consistently smaller compared to the other memory processes for targets, but is the same speed or faster than the other memory processes. Likewise, S_2 is generally greater than the other memory processes for related distractors, but is the same speed or slower than the other memory processes. The latency parameters and core parameters are measuring different things: one measures relative speed, while the other measures strength.

CHAPTER 6

EXPERIMENT 2 AND EXPERIMENT 3

GENERAL DISCUSSION

Classic ideas surrounding dual-process accounts have an assumption that one of the defining differences between familiarity and recollection is that the former is faster than the latter. However, studies in the RK literature as well as recognition studies with imposed deadlines and manipulations that affect recollection have failed to support this hypothesis (Gillund & Shiffrin, 1984; Dewhurst et al., 2006). Despite this, the idea that familiarity is faster than recollection is appealing because it can explain why an inverted U function is observed for false alarms over time in recognition. Fast familiarity, which supports false memory, produces the left arm of the curve, while a mixture of familiarity and slower recollection produces the right arm of the curve.

While this explanation is logically sound, there are several weaknesses. First, prior methods do not directly measure process speeds, and thus conclusions made from them are inferences at best. Second, prior ideas regarding familiarity and recollection fall out of univariate accounts of recollection. Recently, it has been suggested that recollection is bivariate with a separate process for target and context recollection (Brainerd et al., 2015). In the case of the inverted U function, this can be explained by fast context recollection (which supports false memory), and slower familiarity and target recollection (which supports both false and true memories). Third, inverted U functions on their own are not able to discriminate process speeds for list items, since they only use related distractors. It may be the case that recollection and familiarity behave differently for list items than for related distractors, but this is impossible to discern using inverted U functions. Last, inverted U functions may also be due to bias processes.

We may expect the same result if bias decreases as response time increases (Goethe & Oberauer, 2008; Rotello & Heit, 1999).

Latency extension models can address most of the weaknesses that have been discussed. Regarding the first point, latency extension models provide parameter outputs of response latencies, allowing them to be compared with statistical tests. For the second point, a latency extension using a bivariate recollection model is simple to implement, and both Experiment 2 and Experiment 3 took advantage of this with the conjoint recognition model. With the extended conjoint recognition model, it was possible to compare relative speeds of familiarity and the two kinds of recollection. Regarding the third point, because the conjoint recognition model provides parameter estimates of retrieval processes for both related and unrelated distractors as well as for list items, a latency extension of the conjoint recognition model can provide estimates of speeds from processes used for all types of cues, not just for related distractors. Last, while the latency model itself could not answer if bias process speeds were responsible for the inverted U function, data from Experiment 3 was able to show that this was not the case.

The main objective of the last two experiments of this dissertation was to show how a new modeling methodology can provide insight into the speed of memory processes that could not previously be measured directly. Before we can use the model to analyze memory processes, we must first ensure that real data could be fit to the model. It did for both Experiments 2 and 3. In Experiment 2, the model fit four out of six datasets used in the experiment, and of the two that didn't fit, fit failure could be attributed to large individual differences in subject responses. When individual differences were accounted for, overall model fits were acceptable for all datasets. In Experiment 3, model fits for the latency extended model were acceptable in all conditions where the base model fit the data.

There were two main goals for Experiment 2, and two main goals for Experiment 3 that followed up on some lingering questions that remained after Experiment 2. For Experiment 2, the first goal was to see if a latency extension model could be applied to the conjoint recognition model, and fit to real data. It could. The second goal was to see if familiarity was faster than both target and context recollection, as per classic assumptions, or if context recollection was faster than familiarity. Results from Experiment 2 showed that the latter was true: context recollection was significantly faster than familiarity.

Experiment 3 addressed some concerns that may have remained after Experiment 2. First, it may be argued that while context recollection was faster than familiarity in an untimed condition, process speeds may be variable when response deadline manipulations are imposed. Therefore, it is possible that the model results from Experiment 2 are not sufficient to argue that context recollection is faster than familiarity under different experimental manipulations. Experiment 3 addressed this concern by implementing response deadlines in a conjoint recognition design, and then estimating process speeds for each deadline condition. Despite the response deadline manipulation, the relative speed of memory processes were stable for all conditions, and consistent with findings from Experiment 2: phantom recollection was still faster than familiarity. Second, parameter estimates of the core retrieval processes should match predictions from estimates of their latencies. For example, if phantom recollection is fast, then values of phantom recollection should be greater at faster deadlines, thus contributing to elevated false alarms at earlier deadlines. This was observed in Experiment 3. I will discuss the findings of the different memory processes from both experiments in the following section.

Memory Processes

Related Distractor Processes

Retrieval processes from related distractors were of particular interest for Experiments 2 and 3 because of prior research on false alarms used to infer process speeds. Results from Experiment 2 challenged assumptions that familiarity was a faster process than recollection, and that recollection was a univariate process. According to the latency extension model, phantom recollection was faster than familiarity in four out of six different datasets, and it was also faster than recollection rejection in five out of six different datasets. Experiment 2 provided an initial evidence that phantom recollection, and not familiarity, was the fast process responsible for elevated false alarms at early deadlines, and also supported the bivariate, not univariate, conception of recollection, with the two different recollections operating at different speeds.

There may be concern that the latency modeling results from Experiment 2 were confined to the specific conditions that were used, and that it cannot be extended to studies using different manipulations, and especially those that manipulate response deadline. Is it not possible that the relative order of retrieval processes changes if participants are timed, because it may change the way they approach the retrieval task? Findings from Experiment 3 did not support this idea of process variance. Experiment 3 replicated the results from Experiment 2, with phantom recollection generally being faster than the other retrieval processes. This finding held true even with experiments using different response deadline conditions. In addition, the results from the latency models were consistent across both experiments despite Experiment 3 using a slightly different base model than Experiment 2, where erroneous recollection rejection was included. This provides evidence that results of the latency extension model do not depend on deadline manipulations, and also that they are not affected by the separate binning procedures used for

each deadline condition. Therefore, we can use the modeling results to identify why the inverted U function is observed for false alarms over time.

The inverted U function for false alarms to related distractors was observed in Experiment 3. If it were the case that fast phantom recollection was the reason for the left arm of the curve, then it should follow that phantom recollection should be elevated at faster deadlines compared to other processes. Indeed, results showed that phantom recollection was greater than the either familiarity or recollection rejection at the fastest deadline. Also as hypothesized before, familiarity and recollection rejection increase over time, suggesting that a mix of the two contributes to the right arm of the inverted U function. Recollection rejection in particular increased above phantom recollection after the 500ms deadline condition, which coincided with the dip in bias corrected levels of false memory. In addition, the possibility that the inverted U function could be due to bias processes decreasing over time could be ruled out, because the curve was still present even after bias was accounted for with correction.

Target Processes

An advantage to latency extension models is that not only do they provide estimates for the relative speed of retrieval processes for related distractors, but they also provide estimates for the relative speed of retrieval processes for targets. For related distractors, it was found that context (phantom) recollection was faster than familiarity and recollection rejection, a form of target recollection. One important question is whether the relative speed of processes for targets is consistent with the relative speed of processes for related distractors.

For targets, there is identity judgment (a mix of target and context recollection), familiarity, and erroneous recollection rejection (context recollection). In both Experiment 2 and Experiment 3, there were no differences in speed between identity judgment and familiarity.

This is consistent with the findings from related distractor retrieval processes: if context recollection is faster than familiarity, and target recollection is not, then a process that represents the mix of the two types of recollection may average to be the same speed as familiarity. In addition, context recollection for targets, *E*, could be measured in Experiment 3. Similar to the behavior of phantom recollection for related distractors, *E* was either faster or the same speed as familiarity in the model analysis. This lends support that *P* and *E* are tapping into the same type of recollection, and that context recollection is faster than familiarity.

Comparisons With Other Methodologies

Experiment 2 and Experiment 3 were able to demonstrate that a simple latency extension to an existing memory model could directly address some questions regarding memory process speeds that could not previously be answered. Unlike prior theories that posited familiarity to be a fast retrieval process and recollection to be a slower, more effortful retrieval process, the results from these two experiments paint a more nuanced picture. Familiarity was the same speed as target recollection, and significantly slower than context recollection. This finding is consistent with findings from the RK literature, where "Know" responses were slower than "Remember" responses. If familiarity were faster than recollection, then the expectation is that "Know" responses are also faster. However, if context recollection were faster than familiarity and target recollection were the same speed, then it fits with the RK findings as "Remember" responses would be based on recollection of both target and contextual cues, and thus R responses will be the same speed or in some cases faster than K responses.

In addition, the present findings can provide an explanation for the results from the study by Gillund and Shiffrin (1884). In their work, they used speeded recognition designs with manipulations that should affect recollection, with the anticipation that these manipulations

would have a greater effect on recollection at greater response intervals due to recollection being slower. The manipulations, however, did not affect recollection. In the present experiments, recollection (I) was the same speed as familiarity (S_1) for targets: the lack of speed differences between recollection and familiarity for targets supports the null findings of recollection manipulations by speed in prior research.

The unique advantage of latency extension models that are absent in current methods is the fact that process speeds can be directly measured, rather than having to be inferred. Even models that also incorporate time course of processes, such as Cox and Shiffrin's (2017) dynamic model of memory, cannot do this. In addition, the findings from Experiment 2 and 3 are consistent with what was predicted by the dynamic model, suggesting that the two models are compatible with each other but also that the latency extended model provides additional information in the form of parameterized process speeds.

For example, the dynamic model does not assume that memory strength affects the temporal order of memory processes, but that they are fixed, and thus manipulations that affect the strength of memory traces do not change process order. The same result was observed using the latency extension model; the order of memory processes did not change across different types of experimental manipulations, even though their strength changed across different response deadline conditions. This is further evidenced by the fact that, for example, the parameter estimate of familiarity for related distractors, S_2, was generally greater than the other processes P and R, but its latency parameter was generally slower than the other two processes.

Another hypothesis made by the dynamic model is that semantic content is accessed quickly compared to other types of memory traces, such as the modality of a word's presentation. This fits with the finding by the latency extension model that phantom recollection,

a process influenced by semantic gist, is faster than the other two processes. With the case of the latency extension model, however, this conclusion could be made with statistically testable parameter estimates, rather than having to rely on inferences from experimental manipulations.

Concluding Comments

Before latency extension models, the relative speed of memory processes could only be inferred from response deadline studies. By providing parameter estimates for process speeds, latency extension models allow researchers to directly measure them and compare them with statistical tests. In the present experiments, I showed that a fast recollective process, and not familiarity, was responsible for the left arm of the inverted U shape curve: this was impossible to disentangle prior to the model being available. This result is also consistent with findings from the RK literature, which conflicted with classic hypotheses that familiarity was faster than recollection. It also provided additional support for bivariate models of recollection, as the speed for context recollection and target recollection were different from each other. Last, like the dynamic model of memory, the latency extension model supports the hypothesis that process speeds are fixed, and provided quantifiable data to demonstrate this. The latency extension model is a model that is easy to implement, and gives researchers a new tool to study the speed of memory processes, something that could not be measured directly until now.

APPENDIX

Experiment 1 Materials

First instructions:

This is a memory experiment that involves 2 parts. In the first part, you will view 2 lists of vocabulary words that will be presented as Power Point slides. To keep the lists separate, the slides for the first list will be yellow and the slides for the second list will be white. Pay close attention to the slides because there will be a memory test later. As each word comes up on the screen, read it silently to yourself. Between each list, you will also complete some simple math problems.

The second part of the experiment is the memory test. You will receive detailed instructions for the memory test when we get to it. However, it is what we call a list identification test because your task will be to identify words that appeared on the first list versus words that appeared on the second list. Do you have any questions?

We are ready to present the 2 lists now.

Second instructions:

You have just seen 2 lists of vocabulary words. In the following test, you will be asked questions about the words you just saw.

On this memory test, you will respond to 224 short questions about these lists. Here is how the test works. Each question begins with a word that may or *may not* be one of the words that you saw on one or both of the 2 lists. The word is followed by 1 of 3 statements:

- I saw it on **List 1**.
- I saw it on **List 2**.
- I saw it on **List 1 OR List 2**
- I saw it on **List 1 AND List 2**.

You respond to each question by telling us whether you think the statement about the word is correct. You either respond "yes" if you think that the statement is correct and "no" if you think that the statement is incorrect. When you are asked if you saw the word on **List 1**, the correct answer is "yes" if it was on the first (yellow) list, and the answer is "no" otherwise. When you are asked if you saw the word on **List 2**, the correct answer is "yes" if it was on the second (white) list, and the answer is "no" otherwise. When you are asked if you saw the word on either **List 1 OR List 2**, the correct answer is "yes" if it was presented in either List, and the answer is "no" if it was never presented. Last, when you are asked if you saw the word on **List 1 AND List 2**, the correct answer is "yes" if it was on *both* the first (yellow) list and the second (white) list, and the answer is "no" otherwise.

The test will be presented on a computer screen, one word at a time. You will see the word and the question, and then you will hear a brief tone to indicate that you can submit a response. When you think that the answer to a question is "yes," press the **left arrow key**. When you think that the answer to a question is "no," press the **right arrow key**. Please try your best to answer as accurately as possible for all questions, even if you are unsure of the answer.

Now, let's review a few examples of test questions so that you are clear about how to answer them. Here are 8 sample questions and the correct answer for each:

1. Screen:	It was on **List 1**.	Answer = Yes.
2. Cloud:	It was on **List 2**.	Answer = Yes.
3. Magician:	It was on **List 1 OR List 2**.	Answer = Yes.
4. Satin:	It was on **List 1 AND List 2**.	Answer = Yes.
5. Company:	It was on **List 1**.	Answer = No.
6. Algebra:	It was on **List 2**.	Answer = No.
7. Lawnmower:	It was on **List 1 OR List 2**.	Answer = No.
8. Pope:	It was on **List 1 AND List 2**.	Answer = No.

The correct answer to Questions 1 through 4 is "yes" because Screen appeared on the List 1 (the yellow list), Cloud appeared on List 2 (the white list), Magician was on List 1 (the yellow list), and Satin was on both lists. The correct answer to Questions 5 through 8 is "no" because Company was on List 2 and not List 1, Algebra was not on either List, Lawnmower was not on either List, and Pope only appeared on List 1 and not both.

If you have any questions about these instructions, please ask them now.

Study List

		List 1	List 2
Buffer		Satin	Satin
		Kite	Kite
		Gasoline	Gasoline
		Pope	Company
		Magician	Cloud
		Screen	Cork
Categorized			
	Clothing	Hat	Shirt
		Socks	Underwear
		Tie	Suit
		Girdle	Vest
	Birds	Eagle	Robin
		Hummingbird	Bluejay
		Chicken	Swallow
		Starling	Oriole
	Color	Yellow	Blue
		Violet	Indigo
		Gray	Scarlet
		Magenta	Maroon
	Animals (4-legged)	Cat	Horse
		Elephant	Tiger
		Sheep	Fox
		Moose	Zebra
	Fruit	Cherry	Apple
		Pineapple	Strawberry
		Lemon	Fig
		Nectarine	Raspberry
	Kitchen Utensil	Knife	Pot
		Spatula	Spoon
		Cup	Plate
		Toaster	Saucer
	Instrument	Drum	Piano

		Saxophone	Flute
		Bass	Horn
		Piccolo	Harmonica
	Body Parts	Leg	Arm
		Stomach	Toe
		Hair	Shoulder
		Lungs	Ankle
Uncategorized		Mallet	Milk
		Snow	Rod
		Wig	Grasshopper
		Clown	Tea
		Chalk	Lamp
		Phone	Balloon
		Snake	Cement
		Flag	Macaroni
		Ink	Quilt
		Pencil	Pipe
		Frost	Clams
		Pliers	Pillow
		Marijuana	Manure
		Roof	Minnow
		Shrimp	Whale
		Missile	Cake
		Gun	Camera
		Potato	Bread
		Necklace	Rocket
		Butter	Casket
		Tail	Broom
		Bandage	Sugar
		Key	Vinegar
		Carnation	Crayons
		Burro	Sun
		Sleigh	Sand
		Cancer	Mountain
		Beach	Skyscraper
		Trapeze	Tornado
		Blanket	Lantern
		Diapers	Tennis
		Wine	Fur

Test List

L_1 Targets

Categorized				Uncategorized		
	High Frequency	L_1?	Leg	High Frequency	L_1?	Pencil
			Knife			Tail
			Socks			Wine
			Violet			Flag
		L_2?	Cat		L_2?	Key
			Eagle			Ink
			Pineapple			Beach
			Saxophone			Blanket
		L_{1U2}?	Drum		L_{1U2}?	Snow
			Cherry			Snake
			Stomach			Phone
			Spatula			Frost
		$L_{1\&2}$?	Yellow		$L_{1\&2}$?	Roof
			Hat			Butter
			Elephant			Cancer
			Hummingbird			Gun
	Low Frequency	L_1?	Cup	Low Frequency	L_1?	Sleigh
			Bass			Missile
			Moose			Chalk
			Magenta			Pliers
		L_2?	Hair		L_2?	Wig
			Lemon			Mallet
			Starling			Trapeze
			Girdle			Bandage
		L_{1U2}?	Sheep		L_{1U2}?	Clown
			Gray			Potato
			Nectarine			Carnation
			Toaster			Diapers
		$L_{1\&2}$?	Chicken		$L_{1\&2}$?	Marijuana
			Tie			Necklace
			Piccolo			Burro
			Lungs			Shrimp

L_2 Targets

Categorized				Uncategorized		
	High Frequency	L_1?	Horse	High Frequency	L_1?	Camera
			Robin			Tennis
			Strawberry			Lamp
			Spoon			Mountain
		L_2?	Blue		L_2?	Sun
			Shirt			Rod
			Flute			Cake
			Toe			Sugar
		L_{1U2}?	Apple		L_{1U2}?	Bread
			Pot			Pillow
			Tiger			Sand
			Indigo			Tea
		$L_{1\&2}$?	Piano		$L_{1\&2}$?	Pipe
			Arm			Milk
			Bluejay			Balloon
			Underwear			Fur

	Low Frequency	L_1?	Suit	Low Frequency	L_1?	Clams
			Plate			Grasshopper
			Vest			Cement
			Maroon			Broom
		L_2?	Fox		L_2?	Macaroni
			Swallow			Casket
			Raspberry			Skyscraper
			Harmonica			Quilt
		L_{1U2}?	Shoulder		L_{1U2}?	Vinegar
			Fig			Manure
			Oriole			Minnow
			Zebra			Lantern
		$L_{1\&2}$?	Scarlet		$L_{1\&2}$?	Rocket
			Horn			Whale
			Saucer			Crayons
			Ankle			Tornado
Related Distractors						
Categorized				Unrelated Categorized		
	High Frequency	L_1?	Banana	High Frequency	L_1?	Iron
			Green			Pine
			Trumpet			Sofa
			Crow			Platinum
		L_2?	Fork		L_2?	Steel
			Foot			Train
			Rat			Truck
			Shorts			Aluminum
		L_{1U2}?	Guitar		L_{1U2}?	Desk
			Dog			Chair
			Purple			Maple
			Ladle			Redwood
		$L_{1\&2}$?	Cardinal		$L_{1\&2}$?	Oak
			Jacket			Car
			Elbow			Couch
			Watermelon			Airplane
	Low Frequency	L_1?	Lavender	Low Frequency	L_1?	Cedar
			Lime			Lead
			Banjo			Helicopter
			Squirrel			Alloy
		L_2?	Harp		L_2?	Boat
			Wolf			Willow
			Strainer			Sycamore
			Lungs			Bookcase
		L_{1U2}?	Duck		L_{1U2}?	Brass
			Stove			Chest
			Vulture			Uranium
			Apricot			Hickory
		$L_{1\&2}$?	Teeth		$L_{1\&2}$?	Cab
			Belt			Footstool
			Chartreuse			Subway
			Blouse			Bureau
				Uncategorized		
				High Frequency	L_1?	Mirror
						Stick
						Magazine
						Straw

	L_2?	Jail
		Paper
		Needle
		Button
	L_{1U2}?	Hammer
		Box
		Rope
		Rice
	$L_{1\&2}$?	Coffin
		Grass
		Mud
		Basket
Low Frequency	L_1?	Movie
		Kerosene
		Hatchet
		Gym
	L_2?	Volcano
		Quill
		Sponge
		Thermometer
	L_{1U2}?	Freckles
		Pickle
		Dentist
		Chisel
	$L_{1\&2}$?	Sidewalk
		Puddle
		Typewriter
		Pedal

Experiment 3 Materials

Study Instructions

The experiment is divided into three independent study-test cycles. In the study part of each cycle, which we will call the 'study phase', you will be presented with several word lists. They will be presented in an automated fashion. As each word comes up on the screen, read it silently to yourself and pay close attention to them, as your memory will be tested later on.

After each study phase, your memory for the word lists will be tested. We will call this the 'test phase'. You will receive detailed instructions about the memory test later on.

Press ‘u’ when you understand the experiment structure.
We will now begin the study phase. When you are ready, press “s” to begin.

Test Instructions V Condition

You will now take a recognition test.
You will be shown a word that may be either:

(a) a word that you saw in the previous study phase
(b) a new word that is related to a word you saw in the previous study phase
(c) a new word that is unrelated to the words you saw in the previous study phase

For each word, your job is to respond ‘yes’ (press the "left arrow key") if the word was presented during the study phase, and ‘no’ (press the “right arrow key”) otherwise. That is, you should respond ‘yes’ if the test item is from category (a) and ‘no’ if the test item is from category (b) or (c).

Press ‘h’ to see an example of this.

For example, suppose that during the study phase you saw the words “shuttle," "rocket," "astronaut," and "blank."
On the test, if you are given the word “shuttle”, which was on the list, you should answer ‘yes’.
However, if you are given the word “space”, which was not on the list but is related to a word that was presented, you should answer ‘no’. Similarly, if you are given the word “whale”, which was not on the list and is unrelated to any word that was presented, you should answer ‘no’.

Press ‘x’ to learn about how much time you have to respond.

Test Instructions G Condition

You will now take a recognition test.
You will be shown a word that may be either:

(a) a word that you saw in the previous study phase
(b) a new word that is related to a word you saw in the previous study phase
(c) a new word that is unrelated to the words you saw in the previous study phase

For each word, your job is to respond ‘yes’ (press the "left arrow key") if the word was not presented but RELATED to a word that was presented during the study phase, and ‘no’ (press the “right arrow key”) otherwise. That is, you should respond ‘yes’ if the test item is from category (b) and ‘no’ if the test item is from category (a) or (c).

Press ‘h’ to see an example of this.

For example, suppose that during the study phase you saw the words “shuttle," "rocket," "astronaut," and "blank."
On the test, if you are given the word “shuttle”, which was on the list, you should answer ‘no’.
However, if you are given the word “space”, which was not on the list but is related to a word that was presented, you should answer ‘yes’. Similarly, if you are given the word “whale”, which was not on the list and is unrelated to any word that was presented, you should answer ‘no’.

Press ‘x’ to learn about how much time you have to respond.

Test Instructions VG Condition

You will now take a recognition test.
You will be shown a word that may be either:

(a) a word that you saw in the previous study phase
(b) a new word that is related to a word you saw in the previous study phase
(c) a new word that is unrelated to the words you saw in the previous study phase

For each word, your job is to respond ‘yes’ (press the "left arrow key") if the word was presented or related to a word that was presented during the study phase, and ‘no’ (press the “right arrow key”) otherwise. That is, you should respond ‘yes’ if the test item is from category (a) or (b) and ‘no’ if the test item is from category (c).

Press ‘h’ to see an example of this.

For example, suppose that during the study phase you saw the words “shuttle," "rocket," "astronaut," and "blank."
On the test, if you are given the word “shuttle”, which was on the list, you should answer ‘yes’.
If you are given the word “space”, which was not on the list but is related to a word that was presented, you should answer ‘yes’. However, if you are given the word “whale”, which was not on the list and is unrelated to any word that was presented, you should answer ‘no’.

Press ‘x’ to learn about how much time you have to respond.

Test Timing Instructions

In addition, each test word will be displayed for LESS THAN 1 SEC., and will then be followed by a tone. As soon as you hear the tone, you should provide a response (‘yes’ or ‘no’) as quickly as possible. For this reason, you should think about your response BEFORE you hear the tone.

Press ‘r’ for information on response feedback.

Whenever you provide a response, your reaction time (RT) will be displayed soon afterwards. If your response was too slow, the RT will appear in red, and you should try to be quicker next time. If your response was not slow, the RT will appear in green, which means that you are providing a response within the appropriate time window.

Press ‘s’ for some more tips on the timing of your response.

You should always try to provide a response as soon as you hear the tone, but not before the tone. Responses made before the tone will not be recorded.

Press ‘v’ to continue.

Now let’s see how this works before you receive the actual memory test. If you have any questions, please ask the researcher now. As before, suppose you were presented with the words “shuttle," "rocket," astronaut," "blank," during the study phase.

Study List (Version 1)

Study 1

Critical Distractor	Targets	Critical Distractor	Targets	Critical Distractor	Targets
City	New York	Mountain	Climber	Chair	Table
	Urban		Hill		Rocking
	Suburb		Climb		Swivel
	County		Molehill		Recliner
	Chicago		Peak		Seat
	State		Valley		Stool
	Capital		Summit		Desk
	Country		Steep		Couch
	Streets		Ski		Sit
	Metropolitan		Bike		Sofa
	Town		Goat		Bench
	Village		Glacier		Sitting
Slow	Fast	Cup	Saucer	Spider	Web
	Snail		Measuring		Tarantula
	Turtle		Mug		Arachnid
	Sluggish		Goblet		Creepy
	Quick		Coaster		Bug
	Molasses		Plastic		Insect
	Lethargic		Tea		Crawl
	Speed		Coffee		Fly
	Delay		Straw		Fright
	Hesitant		Handle		Poison
	Cautious		Stein		Bite
	Traffic		Drink		Animal
Cold	Hot	Soft	Hard	Man	Woman
	Shiver		Loud		Lady
	Arctic		Tender		Handsome
	Frigid		Fluffy		Male
	Freeze		Pillow		Person
	Chilly		Downy		Suit
	Frost		Plush		Uncle
	Ice		Cotton		Beard
	Warm		Skin		Muscle
	Winter		Touch		Father
	Snow		Fur		Old
	Heat		Furry		Strong
Pen	Quill				
	Pencil				
	Bic				
	Marker				
	Write				
	Fountain				
	Felt				
	Scribble				
	Cross				
	Leak				
	Crayon				
	Tip				

Study 2

Critical Distractor	Targets	Critical Distractor	Targets	Critical Distractor	Targets
Black	White	Fruit	Kiwi	Sweet	Honey

	Gray		Citrus		Bitter
	Brown		Pear		Sugar
	Coal		Berry		Sour
	Dark		Vegetable		Candy
	Color		Banana		Tart
	Funeral		Orange		Chocolate
	Blue		Cherry		Nice
	Charred		Apple		Taste
	Ink		Ripe		Cake
	Death		Basket		Tooth
	Cat		Juice		Good
Music	Band	Rough	Sandpaper	Smoke	Cigar
	Concert		Smooth		Cigarette
	Jazz		Coarse		Pipe
	Symphony		Tough		Tobacco
	Orchestra		Rugged		Puff
	Rhythm		Bumpy		Chimney
	Radio		Jagged		Lungs
	Melody		Riders		Pollution
	Piano		Uneven		Billows
	Sound		Ready		Ashes
	Instrument		Sand		Fire
	Note		Boards		Blaze
Trash	Garbage	Car	Vehicle	Window	Pane
	Rubbish		Automobile		Sill
	Debris		Garage		Shutter
	Dump		Sedan		Curtain
	Litter		Drive		Door
	Landfill		Van		Ledge
	Junk		Keys		Glass
	Waste		Ford		View
	Sewage		Truck		Screen
	Pile		Bus		Open
	Scraps		Jeep		Frame
	Refuse		Taxi		Breeze
King	Throne				
	Queen				
	Crown				
	Reign				
	Monarch				
	Royal				
	Palace				
	Prince				
	Chess				
	Leader				
	Dictator				
	George				

Study 3

Critical Distractor	Targets	Critical Distractor	Targets	Critical Distractor	Targets
Anger	Rage	Sleep	Nap	Shirt	Blouse
	Mad		Doze		Sleeves
	Enrage		Awake		Collar
	Fury		Drowsy		Shorts
	Temper		Snooze		Button
	Ire		Slumber		Pants
	Wrath		Tired		Polo

	Mean		Rest		Jersey
	Hatred		Snore		Vest
	Fight		Wake		Cuffs
	Hate		Yawn		Tie
	Fear		Blanket		Pocket
Smell	Aroma	Lion	Roar	Needle	Thread
	Scent		Tamer		Syringe
	Whiff		Tiger		Haystack
	Stench		Mane		Injection
	Reek		Fierce		Pin
	Sniff		Den		Thimble
	Perfume		Cub		Sewing
	Fragrance		Cage		Knitting
	Nose		Bears		Prick
	Rose		Jungle		Sharp
	Salts		Pride		Thorn
	Breathe		Africa		Point
Doctor	Physician	River	Mississippi	Bread	Rye
	Nurse		Creek		Loaf
	Stethoscope		Stream		Butter
	Surgeon		Flow		Toast
	Patient		Bridge		Dough
	Clinic		Brook		Crust
	Dentist		Lake		Flour
	Medicine		Barge		Sandwich
	Lawyer		Water		Jam
	Health		Boat		Jelly
	Sick		Swim		Slice
	Cure		Run		Milk
Foot	Toe				
	Inch				
	Ankle				
	Shoe				
	Sandals				
	Sock				
	Hand				
	Boot				
	Yard				
	Kick				
	Knee				
	Walk				

Test List (Version 1)

Test 1

Targets	Deadline (ms)	Related Distractors	Deadline (ms)	Unrelated Distractors	Deadline (ms)
Sluggish	100	Chair	100	Deer	100
Lady	300	Pen	100	Long	300
Climb	300	Soft	100	Butterfly	300
Urban	500	Cup	300	Lose	300
Stool	750	Cold	300	Glasses	500
Marker	750	Spider	300	Spaghetti	500
Loud	750	Slow	500	Girl	500
Measuring	1000	Man	750	Terrific	750
Ice	1000	Mountain	750	Straight	1000

Creepy	1000	City	1000	Close	1000
Test 2					
Targets	Deadline (ms)	Related Distractors	Deadline (ms)	Unrelated Distractors	Deadline (ms)
Reign	100	Black	100	Drug	100
Symphony	100	Fruit	100	Factory	100
Shutter	100	King	500	Trouble	300
Keys	300	Window	500	Shy	300
Rubbish	300	Music	500	Wish	750
Tobacco	300	Car	750	Whiskey	750
Candy	500	Smoke	750	Mutton	750
Coarse	500	Trash	750	High	1000
Grey	750	Rough	1000	Atom	1000
Pear	750	Sweet	1000	Church	1000
Test 3					
Targets	Deadline (ms)	Related Distractors	Deadline (ms)	Unrelated Distractors	Deadline (ms)
Toast	100	Smell	100	Justice	100
Wake	100	Lion	300	Plan	100
Shoe	300	Needle	300	Portrait	100
Mad	500	Shirt	300	Rubber	300
Patient	500	Bread	500	Cabbage	500
Stream	500	Sleep	500	Ghost	500
Fragrance	750	Foot	750	Lamp	500
Injection	1000	Anger	1000	Command	750
Sleeves	1000	Doctor	1000	Thief	750
Tiger	1000	River	1000	Stove	1000

REFERENCES

Arndt, J. (2012). The influence of forward and backward associative strength on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 747-756. doi: 10.1037/a0026375

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy in word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–611). New York: Academic Press.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Learning, memory, and thinking* (Vol. 1, pp. 243–293). San Francisco, CA: Freeman.

Ball, B. H., DeWitt, M. R., Knight, J. B., Hicks, J. L. (2014). Encoding and retrieval processes involved in the access of source information in the absence of item memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1271-1286. doi: 10.1037/a0037204

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97,* 548-564. doi: 10.1037/0033-295X.97.4.548

Brainerd, C. J., Gomes, C. F. A., & Moran, R. (2014). The two recollections. *Psychological Review, 121,* 563–599. doi: 10.1037/a0037668

Brainerd, C. J., Gomes, C. F. A., & Nakamura, K. (2015). Dual recollection in episodic memory. *Journal of Experimental Psychology-General, 144,* 816-843. doi: 10.1037/xge0000084

Brainerd, C. J., Holliday, R. E., Nakamura, K., & Reyna, V. F. (2014). Conjunction illusions and conjunction fallacies in episodic memory. *Journal of Experimental Psychology-Learning*

Memory and Cognition, 40, 1610-1623. doi: 10.1037/xlm0000017

Brainerd, C. J., Nakamura, K., Lee, W-F. A. (in press). Recollection is fast and slow. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi: 10.1037/xlm0000588

Brainerd, C. J., Nakamura, K., Reyna, V. F., & Holliday, R. E. (2017). Overdistribution illusions: Categorical judgments produce them, confidence ratings reduce them. *Journal of Experimental Psychology: General, 146,* 20-40. doi: 10.1037/xge0000242

Brainerd, C. J., & Reyna, V. F. (2008). Episodic over-distribution: A signature effect of familiarity without recollection. *Journal of Memory and Language, 58*(3), 765-786. doi:10.1016/j.jml.2007.08.006

Brainerd, C. J., Reyna, V. F., & Aydin, C. (2010). Remembering in contradictory minds: Disjunction fallacies in episodic memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 36,* 711-735. doi:10.1037/a0018995

Brainerd, C. J., Reyna, V. F., Holliday, R. E., & Nakamura, K. (2012). Over-distribution in source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(2), 413-439. doi: 10.1037/a0025645.

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2002). Recollection rejection: false-memory editing in children and adults. *Psychological Review, 110,* 762-784. doi: 10.1037/0033-295X.110.4.762

Brainerd, C. J., & Wright, R. (2005). Forward association, backward association, and the false memory illusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 554-567. doi: 10.1037/0278-7393.31.3.554

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, 27, 307-327. doi: 10.1037///0278-7393.27.2.307

Ceci, S. J., Fitneva, S. A., & Williams, W. M. (2010). Representational constraints on the development of memory and metamemory: A developmental-representational theory. *Psychological Review, 117,* 464–495. doi:10.1037/a0019067

Chen, X. R., Gomes, C. F. A., Brainerd, C. J. (in press). Explaining recollection without remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi: 10.1037/xlm0000559

Cox, G. E., & Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review, 124,* 795-860. doi: 10.1037/rev0000076

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17-22. doi: 10.1037/h0046671

Dewhurst, S. A., Conway, M. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1088-1098. doi: 10.1037/0278-7393.20.5.1088

Dewhurst, S. A., Holmes, S. J., Brandt, K. R., Dean, G. M. (2006). Measuring the speed of the conscious components of recognition memory: Remembering is faster than knowing. *Conscious Cognition, 15,* 147-162. doi: 10.1016/j.concog.2005.05.002

Dosher, B. A. (1984). Degree of learning and retrieval speed: Study time and multiple exposures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 541-574. doi: 10. 1037/0278-7393.10.4.541

Fisher, G. (1997). The jury's rise as lie detector. *The Yale Law Journal, 107,* 575-713. doi: 10.2307/797252

Flegal, K. E., Atkins, A. S., Reuter-Lorenz, P. A. (2010). False memories seconds later: the rapid

and compelling onset of illusory recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1331-1338. doi: 10.1037/a0019903

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16,* 309-313. doi: 10.3758/BF03197041

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67. doi: 10.1037/0033-295X.91.1.1

Goethe, K., & Oberauer, K. (2008). The integration of familiarity and recollection information in short-term recognition: modeling speed-accuracy trade-off functions. Psychological Research, 72, 289-303. doi: 10.1007/s00426-007-0111-9

Heck, D. W., & Erdfelder, E., (2016). Extending multinomial processing tree models to measure the relative speed of cognitive processes. *Psychonomic Bulletin & Review, 23,* 1440-1465. doi: 10.3758/s13423-016-1025-6

Heck, D. W., & Erdfelder, E., (2017). Linking process and measurement models of recognition-based decisions. *Psychological Review, 124,* 442-471. doi: 10.1037/rev0000063

Innocence Project. (2016). DNA exonerations in the United States. Retrieved May 13, 2018. https://www.innocenceproject.org/dna-exonerations-in-the-united-states/

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541. doi: 10.1016/0749-596X(91)90025-F

Jacoby, L. L., Toth, J. P., Yonelinas, A. P., & Debner, J. A. (1994). The relationship between conscious and unconscious influences: Independence or redundancy? *Journal of Experimental Psychology: General, 123,* 216-219. doi: 10.1037/0096-3445.123.2.216

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review,*

87, 252–271. doi: 10.1037/0033-295X.87.3.252

Nakamura, K., & Brainerd, C. J. (2016). Disjunction and conjunction fallacies in episodic memory. *Memory.* doi: 10.1080/09658211.2016.1247869

Pratte, M. S., Rouder, J. N. (2011). Hierarchical single- and dual-process models of recognition memory. *Journal of Mathematical Psychology, 55,* 36-46. doi: 10.1016/j.jmp.2010.08.007

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review, 31,* 180-206. doi:10.1016/j.dr.2011.07.004

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803-814. doi: 10.1037/0278-7393.21.4.803

Roediger, H. L., Watson, J. M., McDermott, K. B., Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8,* 385-407. doi: 10.3758/BF03196177

Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition, 28,* 907–922. doi: 10.3758/BF03209339

Smith, J. B., & Batchelder, W. H. (2008). Assessing individual differences in categorical data. *Psychonomic Bulletin & Review, 15,* 713-731. doi: 10.3758/PBR.15.4.713

Stadler, M. A., Roediger, H. L., McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory and Cognition, 27,* 494-500. doi: 10.3758/BF03211543

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hilssdale, NJ: Erlbaum.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne, 26,* 1-12. doi: 10.1037/h0080017

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293-315. doi:10.1037/0033-295X.90.4.293

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review, 101,* 547-567.

Uyeda, K. M., Mandler, G. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods & Instrumentation, 12*(6), 587-595. doi: 10.3758/BF03201848