The interaction between retrieval and encoding processes in memory

by

Stacey Danckert

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Psychology

Waterloo, Ontario, Canada, 2011

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In memory, encoding and retrieval are often conceived of as two separate processes. However, there is substantial evidence to suggest that this view is wrong—that they are instead highly interdependent processes. One recent example is from Jacoby, Shimizu, Daniels, and Rhodes (2005a), who showed that new words presented as foils among a list of old words that had been deeply encoded were themselves subsequently better recognized than were new words presented as foils among a list of old words that had been shallowly encoded. This paradigm, referred to as memory-for-foils, not only demonstrates a link between encoding and retrieval, but also has led to a proposal about what form this interaction is taking in this task. Jacoby et al. (2005a) proposed that people put in place a retrieval mode that leads to a reprocessing of the original encoding state, which is incidentally applied across both old and new items within the context of a recognition memory test. Such a constrained-retrieval account suggests an intimate relation between encoding and retrieval processes that allows for memories to be highly integrated.

The goal of this thesis is to provide a better understanding of the generalizability and limitations of this memory-for-foils phenomenon and, ultimately, to provide more direct evidence for the interaction of these processes. Experiments 1 and 2 began by replicating the memory-for-foils phenomenon as well as an experiment by Marsh et al. (2009b) which confirmed that the phenomenon does not result simply from strength of encoding differences. Experiment 3 then substituted a deep vs shallow imagery manipulation for the levels-of-processing manipulation, demonstrating that the effect is robust and that it generalizes, also occurring with a different type of encoding. Experiment 4 extended the generalizability of the task to factual phrases. Experiment 5 then moved on to testing the encoding/retrieval interactions by once again employing the imagery encoding manipulation with an additional quality judgment in the final recognition memory test. Using the remember/know paradigm (Gardiner, 1988; Tulving, 1985) demonstrated that more highly-detailed memories were associated with foils from the test of deep items than with foils from the test of shallow items.

From there, response time was used to infer processing speed in Experiment 6a, in a test of whether foils tested among deep items incur an advantage independent of the manipulation undergone by those items. When a lexical decision test replaced the final recognition test, there was no evidence of a memory advantage for “deep” foils over “shallow” foils. Finally, Experiment 6b provided compelling evidence for context-related encoding during tests of deeply encoded words, showing enhanced priming for foils presented among deeply encoded targets when participants made the same deep encoding judgments on those items as were made on the targets during study.

Taken together, these findings provide support for the source-constrained retrieval hypothesis and for the idea of a retrieval mode. New information—information that we may not even be intending to remember— is influenced by how surrounding items are encoded and retrieved, as long as the surrounding items recruit a coherent mode of processing. This demonstrates a clear need to consider encoding and retrieval as highly interactive processes and to avoid conceptualizing them as entirely separate entities. This is a crucial part of increasing our understanding of the fundamental processes in memory.

Acknowledgements

This thesis has truly been a collaborative effort. I am deeply appreciative of the guidance and support offered to me by my supervisors, Dr. Colin MacLeod and Dr. Myra Fernandes. I also offer my thanks to the members of both the MacLeod and Fernandes lab who were available for discussions and assistance along the way. I would also like to thank research assistants Merrick Levine and Molly Pottruff for their invaluable help both in setting up and in running the experiments. Finally, thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Graduate Scholarship (OGS) program for funding my graduate work.

Dedication

I have been tremendously lucky to have had the love and support of my husband along this (very) long journey. Thank you, James, for listening to all of my excitement and frustrations along the way. Thanks for always being my equal and respecting partner. You have been so understanding: knowing when to offer help, when to listen, and when I just needed a hug. And of course, thanks for the endless laughs. I love you always. La vita è bella.

I am also very lucky to have two extremely beautiful boys. You have always made sure to keep me balanced and to give me a smile or hug at exactly the right moment! You have given me countless memories that I need only bring to mind to make any frown turn into a smile. Thank you, Elliott and Cary, for keeping life fun. I love you both more than the whole world.

I am also fortunate to have a wonderful family (both here and far away) as well as many dear friends, who all challenge me and help make life a great place to be. We will definitely be celebrating the completion of this journey! Grandma, you don’t have to ask me anymore: “How much longer until you’re done school?”, I think I am finally done! (Though it may be a little longer before I can definitively answer the question “What are you going to be when you grow up?”!)

Table of Contents

List of Figures

List of Tables

List of Abbreviations

fMRI: functional magnetic resonance imaging

ERP: event-related potential

PET: positron emission tomography

HIPER model: hippocampal encoding/retrieval model

HERA: hemispheric encoding/retrieval asymmetry

CHAPTER 1: INTRODUCTION

> "…the art of remembering is the art of *thinking;* and by adding… that, when we wish to fix a new thing in either our own mind or a pupil's, our conscious effort should not be so much to *impress* and *retain* it as to *connect* it with something else already there. The connecting is the thinking; and, if we attend clearly to the connection, the connected thing will certainly be likely to remain within recall."
>
> William James, Talks to Teachers, 1890

1.1 Background

Since the earliest research on human memory, investigators (Ebbinghaus, 1885) and theoreticians (James, 1890), have sought to categorize memory processes into modular subsystems. Breaking processes into their subcomponents can make the processes simpler to study and understand. However, doing so also leads to a decreased emphasis on how various processes work together and may even create illusions of distinctions that do not actually exist. One potential false distinction is that between encoding (the laying down of new memories) and retrieval (the bringing of past experiences to mind) processes. As Hintzman (2011) rightly notes in criticizing how we have studied memory, there has been a substantial emphasis on the dissociation between these two memory-based processes as being utterly separable when it is clear that they are not.

Evidence for the separability of encoding and retrieval systems comes largely from neuroimaging work (e.g., Gabrieli, Brewer, Desmond, & Glover, 1997; Lepage, Habib, & Tulving, 1998). Early work by Gabrieli et al. (1997) used fMRI to compare the patterns of activation for words of previously viewed outlines of objects and ones that had not been seen previously. They found a region in the anterior medial temporal lobe (i.e., the subiculum portion of the hippocampus) to be related to the retrieval of old items, whereas a posterior portion of the medial temporal lobe (i.e., the parahippocampal gyrus) to be more related to encoding of new words. Based on similar findings using positron emission tomography (PET), Lepage and colleagues (1998) termed this pattern of activation the HIPER model (hippocampal encoding/retrieval model). Other work focusing on more anterior portions of the brain lead to the conclusion of a lateralization of encoding and retrieval processes (Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). Based on this research, Tulving et al. (1994) postulated the HERA model of memory (hemispheric encoding/retrieval asymmetry), in which the left prefrontal cortex (PFC) was associated with encoding while the right PFC was associated with retrieval (see also Desgranges, Baron, & Eustache, 1998; Düzel, Cabeza, Picton, Yonelinas, Scheich, Heinze, & Tulving, 1999). These theories were built from the conceptualization that encoding and retrieval are separate processes. However, modularity cannot be assumed based on evidence of separate brain regions for each "task"; indeed, the regions could work within a network at differing levels for each task where one region may play a larger role than the other in their respective tasks but both are still necessary. The sheer number of requirements of the brain would, perhaps, necessitate all process overlap that is possible, given the limited number of structures available. Therefore, it seems unlikely that, since they require access to much of the same information that encoding and retrieval processes would not overlap.

Researchers investigating amnesia or aging effects on memory often cite either encoding or retrieval as being the principle underlying cause of the memory impairments (e.g., Cipolotti, & Bird, 2006; Dannhauser, Shergill, Stevens, Lee, Seal, Walker, & Walker, 2008; Mark & Rugg, 1998; Van Damme, & d'Ydewalle 2009; Ward 2003). One example of this comes from a neuroimaging study in which Ward (2003) was able to demonstrate some encoding difficulties in patients with frontal lobe damage by looking at new information acquired since injury as compared to information from before the injury. Thus, linking a separate encoding mechanism with the frontal lobe. In contrast, deficits in Korsakoff's syndrome and mild cognitive impairment (thought to be a precursor to Alzheimer's Disease) have been linked to problems in retrieval processes (Van Damme & d'Ydewalle, 2009; Bai, Zhang, Watson, Yu, Shi, Yuan, Zang, Zhu, & Qian, 2009). A study focusing instead on encoding and retrieval during normal aging using ERP suggested a slowing in neural responses during retrieval (Mark & Rugg, 1998) whereas others point to deficits in encoding processes (Cansino, Trejo-Morales, & Hernández-Ramos, 2010). While these studies suggest separate encoding and retrieval processes, it is impossible to find tasks that isolate a single process (see Jacoby, 1991). The 'remaining' process is still likely affected in some way, even if it may be hard to measure.

It is important not to investigate encoding and retrieval with the conception that they are independent, despite the fact that they can appear to be at times. There are hints of their interaction within something as intuitively simple as the levels-of-processing framework, in which deeply encoded information (i.e., semantically encoded information) is more successfully retrieved on a subsequent test than is shallowly encoded information (i.e., perceptually-based encoding; Craik & Lockhart, 1972). Semantic encoding, by its very nature, requires that information be integrated into personal concepts, which requires retrieval and re-storing of past information to succeed (i.e., Is the item pleasant or unpleasant?). In contrast, shallow, surface-based, encoding requires no such links and allows for only very weak and short-lasting memories of the items (i.e., Is there a letter 'a' in the word or not?). If this concept is pushed even further by encoding items in a self-referential way, they are even better remembered later on—having been completely integrated into the pre-existing framework of memory (Rogers, Kuiper, & Kirker, 1977).

While this analysis provides only a hint of how intertwined encoding and retrieval processes are, the transfer appropriate processing principle clearly demonstrates their intimate connection (see Morris, Bransford, & Franks, 1977; Roediger, 1990). This principle states that memory retrieval will be best when the processes invoked during retrieval match those undergone during encoding. In transfer appropriate processing, even items undergoing poorer encoding manipulations are advantaged when a similar task is used at retrieval (Morris et al., 1977). Clearly, then, the encoding and retrieval processes are highly interdependent.

This is especially "front and centre" in the proceduralist perspective proposed by Kolers (Kolers, 1973; Kolers & Roediger, 1984; for a recent review, see Roediger, Gallo, & Geraci, 2002). Kolers argued that distinguishing encoding from retrieval is not really meaningful because every encoding event is a retrieval event and every retrieval event is an encoding event. The operations applied during a nominal study event, or first encounter with a stimulus situation, are reapplied during a nominal retrieval event, or subsequent encounter with—or remembering of—that situation.

Consider a simple everyday illustration. If we are at a grocery store trying to recall a grocery list, along with items from the actual list we also retrieve other information, such as what we plan to cook for dinner. We are also concurrently encoding new information, such as the layout of the store or what is on sale. Later on, realizing that we forgot to buy something, we remember being at the store, and imagine the aisle where the forgotten item was located. It would be unreasonable to label these episodes as purely encoding events or purely retrieval events. Both involve encoding *and* retrieval. Encoding is simply a convenient shorthand for the first processing of an event; retrieval is the shorthand for that event's subsequent processing. We clearly need to understand how encoding and retrieval align as opposed to focusing only on how they differ from one another. This goal is at the core of this thesis.

1.2 The basic design

Building a better conceptualization of how these processes work together, has, however, been difficult until recently. Six years ago, though, Jacoby, Shimizu, Daniels, and Rhodes (2005a) reported the development of a clever method for examining this very issue (see also Jacoby, Shimizu, Velanova, & Rhodes, 2005b; Marsh, Thadeus Meeks, Cook, Clark-Foos, Hicks, & Brewer, 2009b). In their study, Jacoby et al. had participants study one set of words deeply by judging the pleasantness of each and another set of words shallowly by determining whether each word contained a particular letter. In the test phase that followed, there were separate recognition tests for these deeply and shallowly encoded items, with each set of studied targets mixed with a set of unstudied foils. So far, apart from blocking the study and test of the deeply and shallowly encoded words, this was essentially a standard levels-of-processing procedure. The novel part was the addition of a second recognition test phase. Here, the foils from the first recognition test phase became the targets, intermixed with a new set of foils.

Unsurprisingly, the standard levels-of-processing memory benefit of deeply encoded items over shallowly encoded items was obtained on the first recognition test phase. The focus of the study was what happened on the new second recognition test. Intriguingly, the novel (foil) items from the test of deeply encoded items were better recognized than their shallow foil counterparts. Put simply, there was a levels-of-processing effect on these former foils, just as if they had been studied under different orienting tasks. This memory benefit was apparent despite the fact that, during the first recognition phase, the type of judgment made on those items (i.e., old or new?) was equivalent to the type of judgment made on those that did not reveal a subsequent memory boost (i.e., the "shallow" foils). Therefore, it could not have been the judgment that was directly applied to those items that led to the memory enhancement; instead, it must have been a difference in how that judgment was made for foils in the test of deep items as compared to those within the test of shallow items.

When one set of items essentially comprises the fillers and that set is only there to permit a test of the others—the targets—we should be able to easily dismiss them. However, that clearly does not happen, at least not in this task. It appears from this task that, during the attempted retrieval of previously encountered information, the context of that information is also accessed in association with any newly encountered information. To explain this, Jacoby et al. (2005a) proposed that, during the initial recognition test, the participants constrained their retrieval such that both old and new items were consequently processed in the same fashion as the old items had been during their encoding. Therefore, in attempting to retrieve information from the past, contextual information is used to draw out that information and new information about the current retrieval context is concurrently added to the information being retrieved. That same contextual information is also being added to the new items during encoding. This results in *new* items from the test of deep items also being more deeply processed on the first test and hence being more easily recognized during the final memory test than are new items from the test of shallow items. Consequently, retrieval is another occurrence of encoding and, in the process, that encoding is linked to information retrieved from the past.

Does this make sense? Why would an apparently “extra” process be performed in addition to the one that is required by a particular task? The answer is likely that both the retrieval of old information and the encoding and integration of new information are routinely performed together. That is, when information is in a retrieval state, it is susceptible and undergoes an updated encoding (Hupbach, Gomez, Hardt, & Nadel, 2007). During a recognition judgment, it appears that old information is being retrieved and new information is being integrated concurrently so that all information being judged is then intimately linked and made more easily accessible (within the same or a similar context) on subsequent occasions. If participants are able to constrain retrieval based on an encoding task (or context) that relates incoming information, whether old or new, to the original encoding context, then any item recently processed in the same way could more easily be judged as being familiar. Additionally, new items should receive that same benefit by being processed in the same manner as the old items. This is in line with a study by Hupbach, Gomez, and Nadel (2009), which demonstrated that not only are old items susceptible during a new encoding event, but it is specifically the memory related to their context (i.e., their source) that is reinstated and thus easily altered. Of course, this can also cause memory intrusions: When old information is available during a new encoding event, that old information can be contextually integrated and later misremembered (Hupbach et al., 2007)! Therefore, these examples illustrate a ‘flaw’ in our system that is highly suggestive of there being heavily intertwined encoding/retrieval processes.

It is appropriate at this juncture to introduce the concept of a *retrieval mode*, an idea that Jacoby et al. (2005a) favoured, and that has been suggested in the past (e.g., Karpicke & Zaromb, 2010; Koriat, 2000; Nyberg et al., 1995; Tulving, 1998). This is the process of actually going back in time to think of the encoding context or to try to reinitiate it somehow. This process is all in an effort to reconstruct past information, and this reconstruction is more accurate when a specific processing context can be used. Bartlett (1932) showed the importance of context in his study of “The war of the ghosts”. After reading a passage, participants were asked to recall the passage as accurately as possible. He found that recall was not only more coherent but also fit better with the participant’s perspective than the folk tale otherwise would. This study clearly demonstrated that (1) memory is a reconstructive process and (2) that the participant’s schema (their personal past) was strongly reflected in the resulting memories. It is during the reconstruction of old information that memories are most susceptible to encoding new contextual information. Within the Jacoby et al. (2005a) paradigm, the idea would be that, during a test, participants invoke a type of processing that corresponds to that used during study. This fits well with the concept of constrained retrieval and is consistent with the proceduralist perspective that encoding and retrieval ordinarily go hand in hand. The term “retrieval mode” will be used throughout this thesis.

1.3 Conceptual Framework

As already outlined, this thesis grew out of an interest in the idea that encoding occurs as a normal component of a retrieval attempt during a retention test. The particular focus is on the procedure developed by Jacoby and his colleagues (2005a; Marsh et al., 2009b, Jacoby et al., 2005b), which will be referred to as the “memory-for-foils” paradigm. Because this paradigm appears to establish such encoding during retrieval, albeit indirectly, it is potentially an important finding (see Hintzman, 2011). As a result, the question of its generalizability and its limitations provided the initial motivation for the research reported here. As this work proceeded, emphasis moved to demonstrating that the benefit shown by distractors initially tested among deeply processed targets was in fact the result of applying the same processing to those distracters as was applied to the targets. This would confirm that encoding was occurring during retrieval on the initial test, and fit with the idea of a retrieval mode being in place.

This thesis began with a careful replication of Jacoby’s paradigm. The phenomenon, although not large, proved to be robust. It did appear that a retrieval mode was engaged during initial testing, and that instituting this mode had observable consequences for new items that occurred in its presence. Put simply, the retrieval mode also provided a corresponding encoding context, one whose influence was readily detectable on a subsequent recognition test.

But could the observed benefit be due to something other than a retrieval mode? One possibility (also considered by Marsh et al., 2009b) is that the stronger items—those that had been deeply processed during study—were easier to recognize on the test than were the weaker items, leaving more resources to be applied to the distracters during the test of the deeply processed items. The first new experiment tested the possibility that the benefit was simply a by-product of having been among strongly encoded items on the initial test. Because the levels-of-processing encoding manipulation essentially confounds strength and mode of retrieval, either or both could be leading to the memory-for-foils benefit found by Jacoby et al. (2005a). To disentangle these contributions, instead of using a levels-of-processing encoding manipulation, some items were repeated during study to create a “strong” encoding condition and these were contrasted with non-repeated study items that constituted a “weak” encoding condition (see also Marsh et al., 2009b).

This procedure allowed for strength of encoding to be manipulated in the absence of any type of retrieval mode that could be re-entered: It seems implausible that repetition alone could establish a mode that could be re-entered. Without re-entering and thereby reapplying the initial study manipulation to distractors on the initial test, there should be no differential memory for foils from the test of strongly encoded items—if constrained retrieval indeed underlies the effect.

Alternatively, if greater strength of the target items at test is enough to provide an advantage to the accompanying foils, then foils from the test of strong items should show the same benefit as those in the levels-of-processing version of the task. In line with the constrained retrieval hypothesis, the prediction was that the common benefit of repetition at study would be observed, but without any benefit on the final recognition test for distractors that had been grouped with stronger items during the initial test. Essentially, there would be no mode to re-enter.

Building from this foundation, the next step was to ascertain the generalizability of the memory-for-foils phenomenon. This required changing the initial encoding manipulation to create a different type of mode (or context) that could be re-entered during the recognition test. The manipulation chosen had previously been used successfully by Hourihan (2008; Hourihan & MacLeod, 2011). It required participants to perform either deep or shallow imagery on sets of words. Deep imagery refers to the standard pictorial imagery task in which participants form mental images of the referents of words (see Paivio, 1971). The novel task is shallow imagery, where participants form mental images of the printed lower case letters in their upper case form. Hourihan (2008) demonstrated that this manipulation results in differential encoding; for the purpose of this thesis, this manipulation also appeared likely to provide modes that could be re-entered. Such re-entry offers the opportunity for the original encoding context to be applied to the foils at test, thereby leading to a boost in their memory on a subsequent test. The expectation was that, if the memory for foils effect is not very narrowly connected only to the levels of processing encoding situation, this imagery manipulation should provide another demonstration of the memory-for-foils phenomenon.

A second way of exploring the generalizability of the memory-for-foils effect would be to vary not the task but the materials. It is possible that the effect is specific to the type of materials used in the test rather than to the memory process itself. In this thesis, this was tested by substituting more complex stimuli—fact-based statements—for the single words in the original experimental method. It is possible that while the effect is found using a simplified laboratory-based test of words, it does not translate to more complex stimuli such as complete sentences; if so, this would certainly limit its generalizability (again, see Hintzman, 2011, for concerns about studying memory phenomena using only isolated words). If the effect does hold for sentences, however, then it is more likely to have implications for everyday memory processes. The expectation was that the retrieval mode benefit would be observed even with these more complex materials, thereby supporting the hypothesis that the memory-for-foils effect is related to the memory process and not restricted to a specific type of materials.

The fundamental question still remaining about this phenomenon is whether direct evidence for mode reinstatement could be found. Thus far, the findings do not demonstrate that it was in fact the original mode being re-enacted that was responsible for the benefit. The thesis addresses this question in two ways. The first way is to assess the quality of memories being attributed to the foils from the tests of deeply and shallowly encoded items. That is, is there any evidence that a greater quality of memory is associated with the "deep" foils in a way that would parallel the items that were actually deeply encoded? To do so, the imagery-based design was expanded to include additional assessments of memory for the foils in the final recognition test. In particular, participants were asked to evaluate the perceived quality of the memories and to indicate whether each memory was one that was associated with additional episodic details ("remember"), whether the memory had no such additional details ("know"), or finally, whether the item was new (Gardiner, 1988; Rajaram, 1993; Tulving, 1985; for a similar design in the context of the memory-for foils paradigm, see Marsh et al., 2009b).

If the benefits to the foils from the test of deeply encoded words stem from constrained retrieval operations during that initial test, in which participants re-enter and therefore reapply the original encoding to all items, "deep" foils should undergo deep processing and, possibly, "shallow" foils should undergo shallow processing on that initial test. The memories for those items undergoing a type of incidental "deep" processing should then also be associated with additional details related to the encoding question (i.e., the pictorial image formed). It is unlikely that any added efficiency can be gained through the use of constrained retrieval for the shallowly encoded items. Regardless, since the shallow imagery task involves the imaging of a limited number of letters, the memory for these items is unlikely to be associated with additional details. Therefore, the prediction is that participants should respond with a greater proportion of "remember" responses to deep foils as compared to shallow foils. Since more "remember" responses mean more detail-based memories for those items, and since the memory judgments for the tests of deep and shallow items were the same, it is likely that those memory details result from processing of items within the original encoding context. This finding would thus provide a stronger bridge between source-constrained retrieval and the memory-for-foils phenomenon.

Evidence for a benefit in the number of “remember” responses when encoding and retrieval context match has been previously shown (Gardiner, 1988; Java, Gregg & Gardiner, 1997; Macken, 2002; Dewhurst & Brandt, 2007). For example, Macken (2002) demonstrated that when context at retrieval (in this case, colour and location of words) was consistent with encoding, the number of “remember” responses increased. This was explained as having occurred due to a cueing of recollection of item-context associations previously formed during encoding (though, see Mulligan & Lozito, 2006). This finding held even for a powerful encoding manipulation such as generation (as compared to reading words): “Remember” responses were increased when generation is used at retrieval (Dewhurst & Brandt, 2007). However, the same was not true for read items, possibly indicating that this effect is only maintained with effortful processing (and therefore not with the “shallow” condition used in this thesis). Consequently, because “remember” responses are increased by using the same encoding and retrieval processes, this lends more credence to the suggestion of reinstatement of encoding processes (for constrained retrieval) that results in the memory-for-foils phenomenon.

Whether re-enacting the original mode underlies the memory-for-foils benefit (i.e., source-constrained retrieval) may be revealed even more clearly through examination of response time data. It is possible that deeply encoded items receive preferential processing over their shallow counterparts in a way that benefits new items associated with them, such as through speeded response times for those new items. Therefore, the second way we tested for a direct connection between the reinstatement of the original encoding manipulation and the deep foil advantage began by attempting to discount the possibility that there is a unilateral benefit to “deep” foils in terms of response times. Toward this end, the final recognition test was replaced with a lexical decision task that included both types of foils as well as novel items (nonwords). No difference was expected in the speed of performing the lexical decision task for the “deep” foils as compared to the “shallow” foils given that lexical decision was unrelated to how those items were initially encoded—in fact, both sets of items should be similarly primed by their prior exposure.

This leads directly to the critical final experiment. The lexical decision task was unrelated to the original depth of processing task at encoding. If the final task was instead closely related to the encoding task used in processing of the original items, then according to the retrieval mode account differential speed of processing should be observed on that final test. What could be more similar to the original encoding task than that very task itself? As such, in the last experiment of this thesis, the final test was altered to be another instantiation of the original encoding orientation (i.e., "Is the item pleasant/unpleasant?" and "Does the word contain an 'a'/no 'a'?"). These final orienting tasks were limited to words that had not received the orienting questions originally—the "deep" and "shallow" foils from the recognition test, mixed with new items.

If, indeed, during the initial test, the distractors undergo the same processing as was applied to the targets during study as an aid to retrieval (i.e., source-constrained retrieval), then participants should be primed to perform the same orienting task on the foils more quickly than on the new items because the previously processed foils have in essence undergone this task before. Therefore, if participants are faster to respond to the questions originally posed concerning targets during the study phase when they are applied to the foils during the final test (in particular those from the test of deep items), this would provide convincing evidence that these items had previously undergone that specific type of processing during their first encounter (i.e., on the initial test). This would constitute strong support for source-constrained retrieval underlying the memory-for-foils effect.

Together, these experiments demonstrate how dramatically retrieval can influence encoding, contradicting a purely separate conceptualization of encoding and retrieval processes. The principal approach used in this thesis involved variations on the experimental design originally created by Jacoby and colleagues (2005a), one that cleverly demonstrated an advantage for foils initially tested among deeply encoded items as compared to those tested among shallow items when they were subsequently tested. Despite only being instructed to retrieve old items from memory, participants concurrently (and incidentally) encoded new items within the same context as those items being retrieved. This clearly points to highly integrated encoding and retrieval processes. The interaction between these processes is maintained over a variety of situations and with different stimuli, making it likely that this type of interaction is one that would be found beyond laboratory settings.

In terms of an efficient memory system, this encoding-retrieval partnership makes a great deal of sense. It is unlikely that one would come across new but completely irrelevant information in the context of another set of old information. Automatically linking any new information to pre-existing information by using the relevant context would make future attempts to access that once new information easier. Without such links, all new information would potentially be new and difficult to access—"islands" of information within memory. Indeed, there is some evidence relating to autobiographical memories in amnesics that this may occur (Medved, & Hirst, 2006). Encoding while retrieving prevents this isolation.

CHAPTER 2: REPLICATIONS

2.1 Experiment 1: Replication of Jacoby et al. (2005a)

Before beginning to address the specific questions about the central finding in the Jacoby, Shimizu, Daniels, and Rhodes (2005a) study, a faithful replication was carried out of their Experiment 1, in which the memory-for-foils effect was first reported. Jacoby, Shimizu, Velanova, and Rhodes (2005b) had previously replicated the phenomenon, as had Marsh et al. (2009b) independently, but it was deemed important to be certain that the same basic pattern was obtained at the outset in this series of studies.

2.1.1 Method

Participants. A total of 24 undergraduate students (20 female, 4 male) from the University of Waterloo participated for credit or remuneration ($5). Their mean age was 20.9 years (SD=1.5).

Materials. The stimuli consisted of 247 words, 5–8 letters long, taken from the MRC Psycholinguistic online database (www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm). The words had an average length of 5.3 letters and an average frequency of 37 occurrences per million (Kučera & Francis, 1967). Words were randomly assigned to six lists of 36 words each, such that each participant had a unique assignment of items. Each phase began and ended with an additional three-word buffer; these “buffer” words were not included in any analyses. In all phases, the words were presented in lowercase letters. Two raters rated approximately half (53%) of the items as being pleasant. Of course, due to the subjective nature of such a rating task, there likely would be high variability in such ratings. Half of the items contained an “a,” and the remaining half of the items did not.

Procedure. A schematic of the experimental procedure is displayed in Figure 2.1. All three phases were participant-paced and were identical to those of Jacoby et al. (2005a). All stimuli were displayed in white lowercase font on a black computer screen. Participants were tested individually and completed the entire experiment in approximately 30 min.

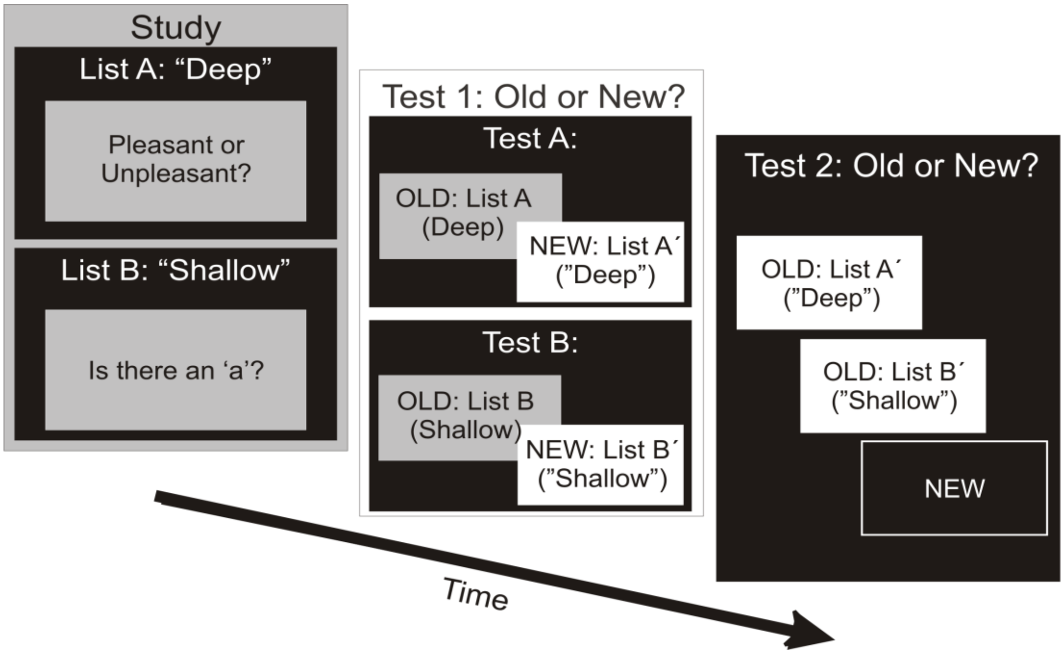

Figure 2.1. Experiment 1: Schematic of the procedure. Every participant performed the two study sessions, then the two corresponding components of Test 1, and then Test 2.

In the study phase, participants performed two encoding tasks, with their order counterbalanced across participants. In the deep encoding task, 36 words were presented one at a time and participants were to indicate, for each, whether it represented something that was pleasant or unpleasant for them. In the shallow encoding task, participants viewed a different 36-word list and were to indicate whether each contained the letter "a", by pressing 1 for "a", 0 for no "a" on the keypad. Following the participant's keypad response to each word, a fixation cross was then presented for 500 ms.

Next came the first recognition phase, Test 1. On two separate 72-item subtests, the 36 deeply encoded words were intermingled with 36 new words, and the 36 shallowly encoded words were intermingled with 36 additional new words. The order of the two subtests was counterbalanced across participants, who were explicitly informed which list the old items were drawn from (e.g., “All old words are from the list for which you made pleasant/unpleasant decisions”). Participants were asked to press 1 for an old item (target) or 0 for a new item (foil) on the numeric keypad.

Finally, there was the second recognition phase, Test 2. Here, the targets were all of the former *foils* from the first recognition phase—from both the deep encoding and shallow encoding recognition tests (i.e., no deeply or shallowly encoded items from the study phase were included on Test 2). Intermixed with these newly defined targets was a completely new set of previously unseen words, such that there were 72 old words (36 “deep” foils and 36 “shallow” foils) and 72 new words. Participants were asked to respond either “old” or “new” by pressing 1 and 0 on the keyboard, respectively.

2.1.2 Results and Discussion

Recognition Test 1. Participants readily distinguished between old and new items from both the deep and shallow encoding tasks (overall hits = .71, overall false alarms = .23). A paired-samples *t*-test was used to compare the recognition scores (hits minus false alarms) for words that had been categorized deeply versus shallowly during study. A typically robust levels-of-processing effect was found, $t(23) = 4.24$, $p < .001$. Figure 2.2a presents the hit and false alarm rates. Separate paired-sample *t*-tests for hits and false alarms revealed a difference only for the hits, $t(23) = 3.76$, $p = .001$ (deep > shallow), and not for the false alarms, $t(23) = 1.06$, $p > .20$. Clearly, participants were effectively using the different encoding techniques during study.

Recognition Test 2. To determine whether there was a levels-of-processing effect for the foils from Test 1 when they became the targets on Test 2, a paired-samples *t*-test was carried out comparing proportions of correct responses across word types. As shown in Figure 2.2b, memory was superior for foils that had initially been tested among deeply encoded words than for foils that had initially been tested among shallowly encoded words, $t(23) = 2.48$, $p < .05$, providing a clear replication of the pattern reported by Jacoby et al. (2005a).

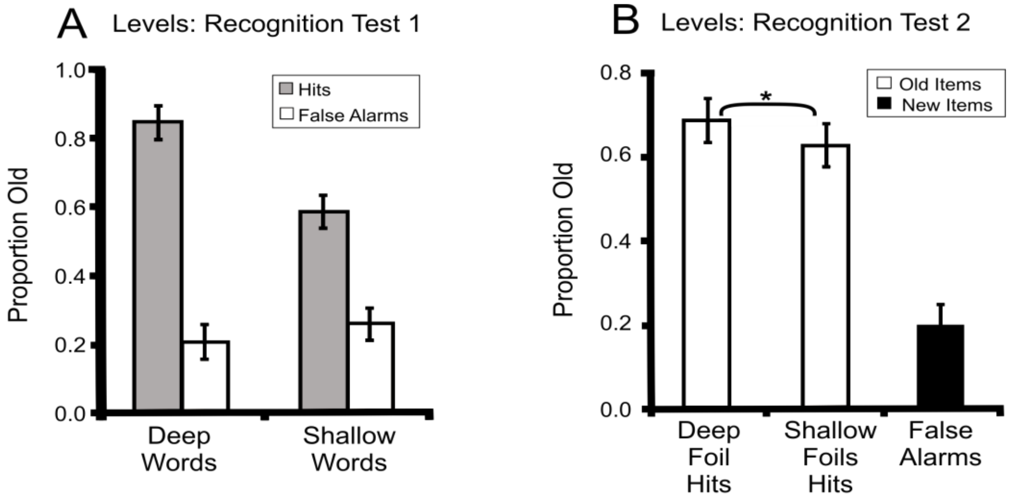

Figure 2.2. Experiment 1: Replication of memory-for-foils effect. (a) Recognition performance on Test 1, demonstrating the levels-of-processing effect. (b) Recognition performance for "old" responses on Test 2, demonstrating the memory-for-foils effect. Together, these data provide a complete replication of Jacoby, Shimizu, Daniels, and Rhodes (2005, Experiment 1). Error bars represent the standard errors of the corresponding means.

Jacoby et al. hypothesized that this memory-for-foils phenomenon was a product of constrained memory searches based on the study context. However, in the original experiment design, the encoding manipulation confounded strength of encoding with mode of encoding, leaving the possibility that the phenomenon could result from the differential strength of encoding of the items. Therefore, the second replication addressed this potential issue.

2.2 Experiment 2: Replication of Marsh et al. (2009b, Experiment 3)

Marsh et al. (2009b, Experiment 3) examined whether the memory-for-foils effect was a by-product of testing words that had been strongly versus weakly encoded during the study phase. In their study phase, words in one list had been studied only once each, whereas those in the other list had been studied three times each. Repetition is well established as influencing strength of encoding (e.g., Hintzman, 1976). Yet it is difficult to imagine how participants could reinstate the number of study presentations as a mode that would constrain retrieval, which means that the memory-for-foils effect should not occur. That is precisely what Marsh et al. observed. Although done independently without knowledge of their experiment, I conducted almost exactly the same experiment, and with the same outcome.

2.2.1 Method

Participants. A total of 24 undergraduate students from the University of Waterloo (13 female, 11 male) participated for credit or remuneration ($5). Their mean age was 21.0 years (SD=3.2).

Materials. The stimulus words were the same as in Experiment 1.

Procedure. All three of the phases were participant-paced. Again, words were randomly assigned to six lists of 36 words each, with different randomizations for each individually tested participant. Three-word primacy and recency buffers were, again, not included in any analyses.

In the study phase, participants studied two 36-word lists—one in which items were each presented once, and one in which items were each presented three times. For the latter, the entire list was randomized three times, with an untested filler word inserted after each of the first two completely randomized list presentations to prevent the unlikely possibility of repeating a word on consecutive trials. For both lists, each word was presented for 1,500 ms, followed by a fixation cross for 500 ms, with no response required. The order of studying the two lists was counterbalanced across participants.

Next came the first recognition phase, Test 1, conducted precisely as in Experiment 1, including instructions to participants concerning the source of the items (e.g., “All old items are from the list in which you saw each item only one time”).

The second recognition phase, Test 2, also was identical to that of Experiment 1, with targets again consisting of all of the former foils from the first recognition phase—from both the “triple” and the “single” recognition tests—among an equal number of new items. Responding was done as on the first recognition test.

2.2.2 Results and Discussion

Recognition Test 1. Participants readily distinguished old from new items for both repeated- and single-presentation lists (overall hits = .74, overall false alarms = .08). The data appear in Figure 2.3a. A paired-samples *t*-test was used to compare the recognition scores (hits – false alarms) for words that had been studied either once or three times. In line with the extensive repetition literature (e.g., Bentin & Moscovitch, 1988; Murnane & Shiffrin, 1991; Ratcliff, Clark, & Shiffrin, 1990; Stretch & Wixted, 1998), participants recognized items that had been studied three times significantly better than items studied only once, $t(23) = 5.76$, $p < .001$. Separate *t*-tests demonstrated that there were (1) more hits, $t(23) = 5.59$, $p < .001$, and (2) fewer false alarms, $t(23) = 2.19$, $p < .05$, for words presented three times than for words presented once.

Recognition Test 2. This time, because repetition appeared unlikely to be a processing mode that could be reinstated, the prediction was that the memory-for-foils effect would not be observed. As shown in Figure 2.3b, there was, in fact, no difference in recognition of former foil words as a function of whether they were encountered on Test 1 among targets presented once or among targets presented three times, $t(23) = 0.84$, $p > .20$. Further, the effect sizes for this test had quite adequate power to correctly reject the null hypothesis (power = .87). This absence of a memory-for-foils effect provides a clear replication of the Marsh et al. (2009b, Experiment 3) pattern and is consistent with the idea that, for new items to be encoded with the same memory benefits as their old counterparts, there must be a coherent mode during the initial encoding session that can then be reapplied to items during a subsequent retrieval session.

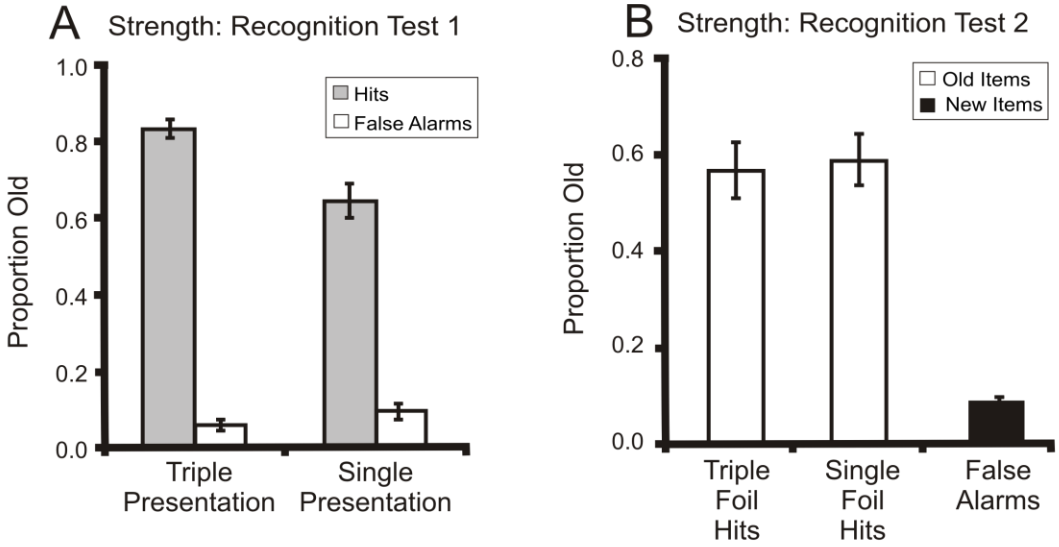

Figure 2.3. Experiment 2: Manipulating strength via repetition at encoding. (a) Recognition performance on Test 1, demonstrating enhanced recognition for words studied three times as opposed to once. (b) Recognition performance for “old” responses on Test 2, showing no evidence of a memory-for-foils effect. Together, these data provide a complete replication of Marsh et al. (2009b, Experiment 3). Error bars represent the standard errors of the corresponding means.

CHAPTER 3: EXTENSION AND GENERALIZATION

3.1 Experiment 3: Imagery

The next experiment was conducted to test the generality of the memory-for-foils effect by substituting a different encoding-retrieval mode for levels-of-processing. To accomplish this, I turned to the venerable encoding task of visual imagery (see Paivio, 1971, 1995, 2007) and to a variant used by Hourihan (2008) in her dissertation. The goal was to create two distinct imagery modes. The first was the standard pictorial imagery task, where participants are instructed to form a mental picture of the word's referent object; this will be referred to as "deep imagery." The second was a case imagery task, where participants are instructed to imagine the presented lower case word all in upper case; this will be referred to as "shallow imagery." Naturally, deep imagery should result in better memory than shallow imagery, as indeed it did in Hourihan's dissertation (2008).

It was expected that these two encoding tasks would form coherent processing modes readily invoked again on the separate subtests forming Test 1. Consequently, the memory-for-foils pattern observed by Jacoby et al. (2005a) and replicated here in Experiment 1 should reappear. This experiment should therefore confirm the robustness and generalizability of the 'memory-for-foils" phenomenon, specifically testing whether levels-of-processing tasks are required during encoding or whether any coherent, reproducible processing mode can also produce the effect.

3.1.1 Method

Participants. Twenty-four undergraduate students from the University of Waterloo (13 female) participated for credit or remuneration ($5). The mean age was 20.28 (SD = 1.5).

Materials. The stimuli consisted of 247 words 5–8 letters in length obtained from the Thorndike and Lorge (1944) norms. The words had an average length of 5.7 letters and an average frequency of 22.1 per million. (Note that these stimuli were from the MRC database, which provided the additional information of word imageability.) All words had moderate to high imageability ratings between 550 and 800. In all phases, the words were presented in lowercase letters. Words were randomly assigned to six lists of 36 words each, with unique randomizations for each participant. In addition, each phase began and ended with three-word "buffers" to discount primacy and recency; these words were not included in any analyses.

Procedure. Participants were tested individually and completed the task in approximately 30 minutes. Apart from the changes in encoding tasks and encoding instructions, all procedures were identical to those of Experiment 1.

In the Study Phase, participants performed a different imagery-based categorization task on each of the two study lists. In the 'deep' imagery task, participants were asked to form a mental picture representing each item. In the 'shallow' imagery task, participants were asked to form a mental image of each word in capital letters (e.g., for cake: CAKE; Figure 3.1, Panel A). Once a participant had created an image, they pressed a key; following this, a fixation cross was presented for 500 ms. The order of performance of the imagery tasks was counterbalanced across participants.

Next came the first recognition phase, Test 1, which was nearly identical to Experiment 1. On two separate 72-item subtests, the 36 deeply imaged words were intermingled with 36 new words, and the 36 shallowly imaged words were intermingled with 36 other new words. The order of the two subtests was counterbalanced across participants, who were explicitly informed which list the old items were drawn from (e.g., "All old words are from the list for which you formed images in your head related to the words"). Participants were asked to press 1 for an old item (target) or 0 for a new item (foil) on the numeric keypad.

Finally, there was the second recognition phase, Test 2. Here, the targets were all of the former foils from the first recognition phase—from both the deep imagery and shallow imagery recognition tests (i.e., no deeply or shallowly encoded items from the study phase were included on Test 2). Intermixed with these newly defined targets was a completely new set of previously unseen words, such that there were 72 old words (36 deep-imagery foils and 36 shallow-imagery foils) and 72 new words. Participants were asked to press 1 for an old item (target) or 0 for a new item (foil) on the numeric keypad.

3.1.2 Results and Discussion

Recognition Test 1. Participants were able to recognize the study lists very well across imagery condition (overall Hits = 0.77, False alarms = 0.08), as shown in Figure 3.1, Panel A. A paired-samples t-test was conducted to compare recognition for words that had been imaged pictorially and those that had been imaged in capitals. Participants had better memory for pictorially imaged as compared to capital-imaged words, $t(23) = 4.31$, $p < 0.001$. This was true for hits, $t(23) = 4.45$, $p < 0.001$, and showed a corresponding trend for false alarms—a greater number of false alarms for capital-imaged words than for pictorially-imaged words, $t(23) = 2.01$, $p = 0.06$ (Figure 3.1, Panel B). Therefore, participants were effectively using the different encoding techniques, resulting in better encoding for words imaged as pictures than for words imaged in upper case.

Recognition Test 2. A paired-samples t-test demonstrated an effect of type of processing, with better memory for old “pictorial” foils than for old “capital” foils, $t(23) = 3.48$, $p < 0.005$. In line with the levels of processing finding in Jacoby et al. (2005a) and in Experiment 1 reported here, there were more hits for foils that had been initially viewed during the test for words imaged pictorially than for foils that had been initially viewed during the test for words imaged as capitals (Figure 3.1, Panel C).

Thus, a different encoding task than the only one that has previously been used also produced the ‘memory for foils’ effect. Participants demonstrated enhanced subsequent recognition for new words tested among words that had been imaged pictorially as compared to new words tested among words that had been imaged in upper case. This deep/shallow imagery manipulation is therefore quite analogous to the deep/shallow levels of processing manipulation. These data demonstrate that using an imagery-based encoding technique provides the same pattern of results as seen in a typical levels of processing study. Clearly, what is important is that the mode of encoding is sufficiently coherent that it can be re-enacted at the time of retrieval. This mode at retrieval then “spills over” onto the foils, producing an encoding benefit for those accompanying items deeply encoded in the preceding

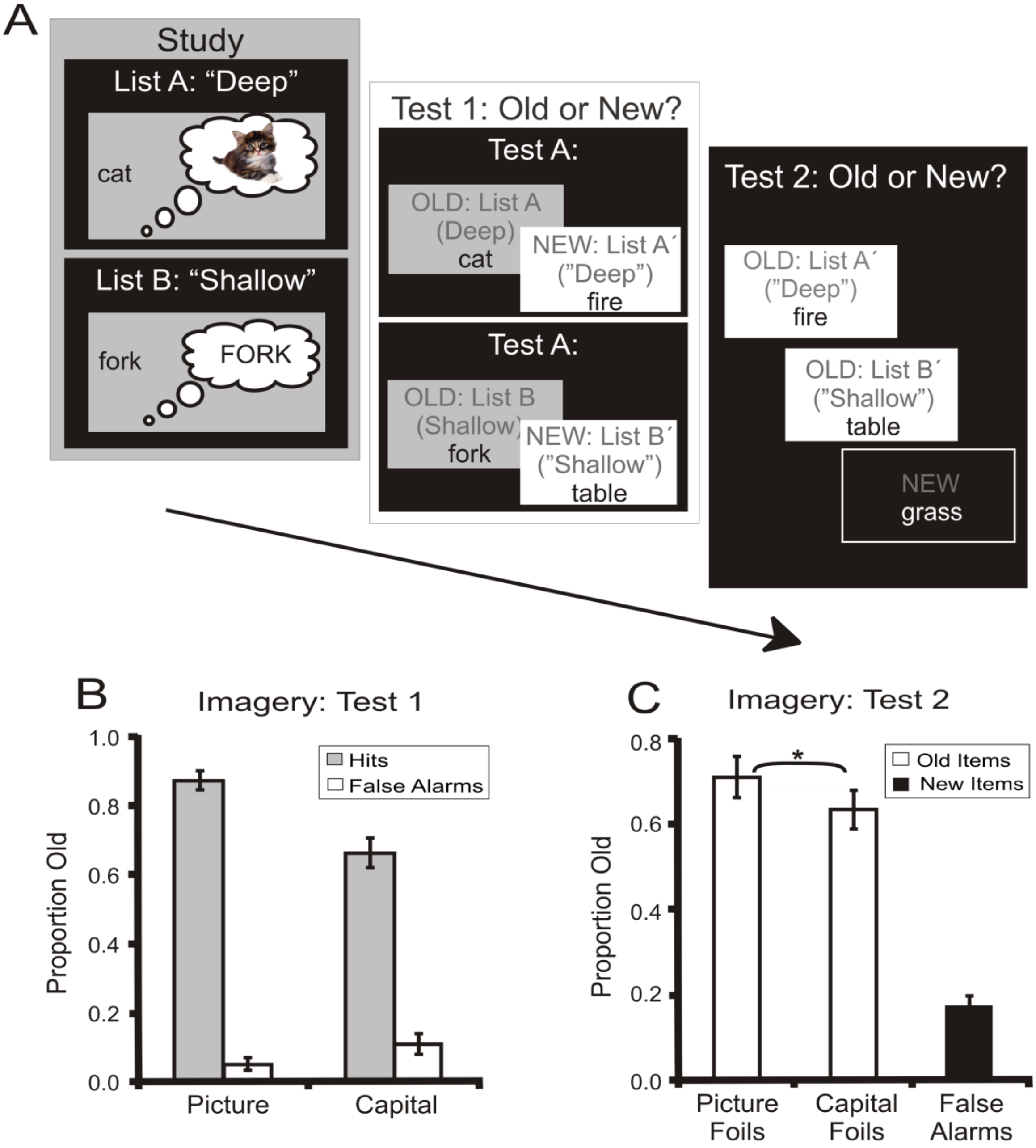

Figure 3.1. Experiment 3: Manipulating type of imagery at encoding. (a) Outline of the procedure. (b) Recognition data from Test 1, demonstrating enhanced memory following pictorial imagery as compared to capital letter imagery. (c) Recognition performance for Test 1 foils on Test 2, demonstrating a clear memory-for-foils effect. Error bars represent the standard errors of the corresponding means.

study phase. This finding supports the source-constrained retrieval hypothesis of Jacoby et al., demonstrating its generalizability.

3.2 Experiment 4: Phrases

The preceding experiment generalized the memory-for-foils paradigm to a different encoding task. To further investigate the generalizability of the paradigm, the materials were altered: Phrases were substituted for the words of the standard Jacoby et al. (2005a) paradigm. This also permitted a test of whether the phenomenon would hold for more complex stimuli and therefore be more applicable beyond the laboratory setting.

3.2.1 Method

Participants. A total of 24 undergraduate students from the University of Waterloo (18 female, 6 male) participated for credit or remuneration ($5). Their mean age was 20.8 years (SD = 2.4).

Materials. The 247 stimulus phrases were gathered from various sources, being selected from internet sources to be similar to those of Gopie and Macleod (2009). The phrases ranged from 3 to 13 words in length. All phrases contained unusual facts. See appendix A for examples.

Procedure. Phrases were randomly assigned to six lists of 36 phrases each, with different randomizations for each individually tested participant. Again, three-phrase primacy and recency buffers were not included in any analyses. Participants were tested individually and completed the entire experiment in approximately 30 min.

In the study phase, participants performed two encoding tasks, with their order counterbalanced across participants. In the deep encoding task, 36 phrases were presented one at a time and participants were to indicate, for each, whether it was believable or unbelievable. In the shallow encoding task, participants viewed a different 36-phrase list and were to indicate whether each contained seven or fewer words or greater than seven words. Participants were given 2,500 ms in which to complete their responses.

The next two phases were identical to Experiment 1 with the exception that, instead of words, phrases were used as stimuli. During the first recognition phase, Test 1, participants completed two separate 72-item subtests: one in which 36 deeply encoded phrases were intermingled with 36 new phrases, and another in which 36 shallowly encoded phrases were intermingled with 36 additional new phrases. The order of the two subtests was counterbalanced across participants, who were explicitly informed which list the old items were drawn from (e.g., “All old words are from the list for which you made believability decisions”). Participants were asked to press 1 for an old item (target) or 0 for a new item (foil) on the numeric keypad.

The final recognition phase, Test 2, was identical to that in Experiment 1 with the substitution of phrases in place of words.

3.2.2 Results and Discussion

Recognition Test 1. Participants readily distinguished between old and new items from both the deep and shallow encoding tasks (overall hits = .65, overall false alarms = .09). A paired-samples *t*-test was used to compare the recognition scores (hits minus false alarms) for phrases that had been categorized deeply versus those that had been categorized shallowly during study. A typically robust levels-of-processing effect was found, $t(23) = 13.83$, $p < .001$. Figure 3.2a presents the hit and false alarm rates. Separate paired-sample *t*-tests for hits and false alarms revealed a difference for the hits, $t(23) = 12.42$, $p < .001$, and also for the false alarms, $t(23) = 2.88$, $p < .005$. Clearly, participants were effectively using the different encoding techniques during study.

Recognition Test 2. To determine whether there was a levels-of-processing effect for the foils from Test 1 when they became the targets on Test 2, a paired-samples *t*-test was conducted comparing proportions of correct responses across word types. As shown in Figure 3.2b, memory was superior for foils that had initially been tested among deeply encoded phrases than for foils initially tested among shallowly encoded phrases, $t(23) = 4.38$, $p < .001$.

Thus, a different type of stimuli than those previously used also produced the memory-for-foils effect. Participants demonstrated enhanced subsequent recognition for new phrases tested among phrases that had been deeply encoded as compared to new phrases previously tested among shallowly encoded phrases. That the memory-for-foils effect is maintained with a more complex type of stimulus clearly demonstrates that the effect is quite robust despite the added complexity. When the stimuli are tested within the context of

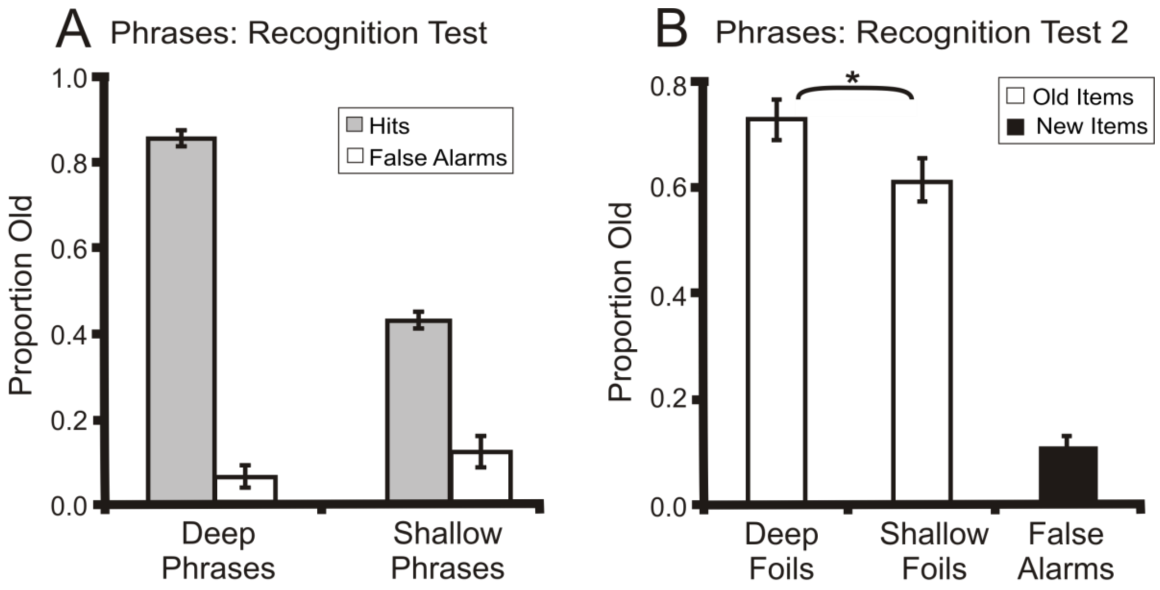

Figure 3.2. Experiment 4: Manipulating type of stimulus at encoding. (a) Recognition data from Test 1, demonstrating enhanced memory following deep encoding of phrases as compared to shallow encoding of phrases. (b) Recognition performance for Test 1 foils on Test 2, demonstrating a clear memory-for-foils effect. Error bars represent the standard errors of the corresponding means.

deep or shallow encoding, the "deep" foils again receive the benefits of the "spill over" from the items that truly did experience a deep encoding. This finding supports the source-constrained retrieval hypothesis of Jacoby et al. (2005a), and additionally demonstrates and expands upon the generalizability of the phenomenon. Further, this provides a basis for potential applicability of this phenomenon.

CHAPTER 4: LINKING ENCODING AND RETRIEVAL

4.1 Experiment 5: Imagery with "Remember/Know" Judgments

After establishing the generalizability of the memory-for-foils effect, a fundamental question remained. Thus far, there has been no direct evidence for mode reinstatement in any of the reported studies (Jacoby et al., 2005a; Jacoby et al., 2005b; Marsh et al., 2009b) or in any of the previous studies in this thesis. Although better recognition of foils that accompany deeply processed targets is consistent with deeper processing of those foils, which in turn is consistent with a deeper mode of processing, that logic is indirect. Therefore, the following studies were designed to address this question.

To start, Experiment 3 of the thesis was modified to include a quality of memory-for-foils so as to further link those memories with a constrained retrieval approach. Stemming from neuropsychological and imaging data (e.g., Aggleton, McMackin, Carpenter, Hornak, Kapur, Halpin, Wiles, Kamel, Brennan, Carton, & Gaffan, 2000; Curran, 2000; Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997, Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998), recognition memory tests actually appear to draw from two separate sources of memory: one that is based on conscious recollection of the encoding event or aspects of it, and the other that is limited to a memory that contains only the familiarity of the old item (Gardiner, 1988). One way to obtain a subjective indication for whether each memory is recollective (i.e., detail-rich) or familiarity-based is through the use of the Remember/Know paradigm (Tulving, 1985; for a review see Yonelinas, 2002). In this paradigm, participants are asked, within the context of a recognition test, to reflect on the qualitative nature of their memories. For detail-rich, recollective memories, they are asked to provide a "remember" response, and for detail-poor memories they are to provide a "know" response.

An analysis of the qualitative nature of the memory-for-foils effect can provide additional support for the constrained retrieval hypothesis. If the initial encoding manipulation is re-instituted on foils during the first recognition test, more “remember” responses would be linked to recognition associated with greater detail, which would be expected for foil words from the deep imagery test (similar to the findings of Marsh et al., 2009b, Experiment 1). For the detail-poor “shallow” foils, there should be fewer “remember” responses. This experiment represents a conceptual replication of Experiment 3 in which a remember/know judgment was included in the final recognition test.

4.1.1 Method

Participants. A total of 25 undergraduate students from the University of Waterloo (21 female, 4 male) participated for credit or remuneration ($5). After one female participant was removed for failing to comply with instructions on the final test, the mean age was 20.3 years (SD=3.3).

Materials. The stimuli were identical to those in Experiment 1.

Procedure. All three phases were participant-paced. Participants were tested individually and completed the entire experiment in approximately 30 min. Testing parameters were identical to those in Experiment 3.

The study phase and the first recognition phase, Test 1, were identical to Experiment 3.

Finally, there was the second recognition phase, Test 2. Here, the targets were all of the former foils from the first recognition phase—from both the deep imagery and shallow imagery recognition tests (i.e., no deeply or shallowly encoded items from the study phase were included on Test 2). Intermixed with these newly defined targets was a completely new set of previously unseen words, such that there were 72 old words (36 deep-imagery foils and 36 shallow-imagery foils) and 72 new words. Participants were asked to respond based on the quality of their memories, saying either “remember.” “know,” or “new.” They were given very careful instruction and practice on deciding whether the words were new, or were old and accompanied by detailed memories (i.e., “remember” response), or were old and not accompanied by any detailed memories (i.e., “know” response). The instructions closely followed those used by Gardiner (1988, p. 311), including use of the examples that they provided.

4.1.2 Results and Discussion

Recognition Test 1. Participants were able to recognize the study lists very well across imagery conditions (overall hits = .78, overall false alarms = .11), as is shown in Figure 4.1a. A paired-samples *t*-test demonstrated that participants had considerably better overall memory for pictorially imaged words than for words imaged in capitals, *t*(23) = 6.74, *p* < .001. This was true for hits, *t*(23) = 7.44, *p* < .001, and showed a complementary pattern for false alarms—more false alarms for capital-imaged than for pictorially imaged words, *t*(23) = 2.19, *p* < .05. Therefore, participants were effectively using the different encoding techniques, resulting in better encoding for words imaged as pictures than for words imaged in uppercase.

Recognition Test 2. Most important, a paired-samples *t*-test demonstrated a significant effect of type of imagery, with better memory for old pictorial foils than for old capital foils, *t*(23) = 3.41, *p* < .005. Thus, the memory-for-foils effect, as generalized to an entirely different form of encoding

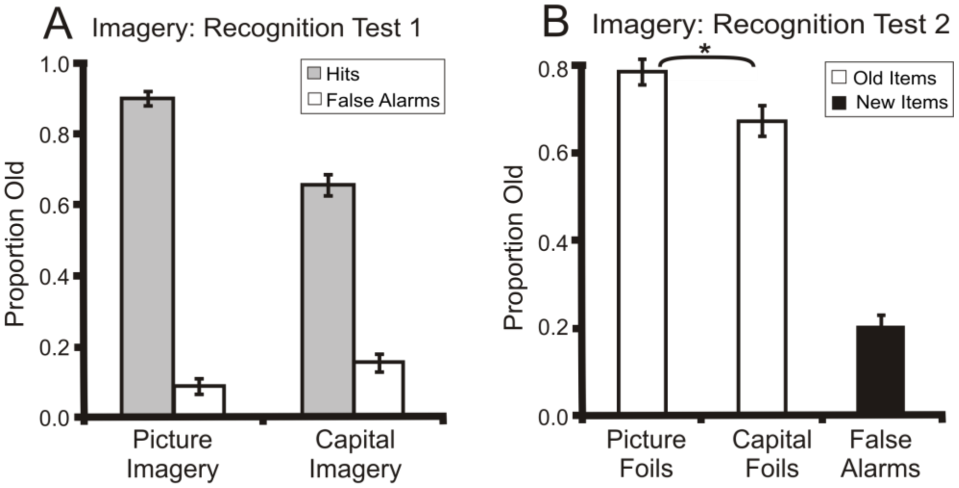

Figure 4.1. Experiment 5. Manipulating type of encoding and inclusion of a quality of memory assessment. (a) Recognition data from Test 1, demonstrating enhanced memory following pictorial imagery as compared to capital-letter imagery. (b) Recognition performance for Test 1 foils on Test 2, demonstrating a clear memory-for-foils effect. Error bars represent the standard errors of the corresponding means.

manipulation, replicated. In line with the levels-of-processing finding in Jacoby et al. (2005a), there were more hits for foils that had initially been pictorially imaged during the first test than for foils that had initially been imaged as capitals (Figure 4.1b). The pattern perfectly replicated that of Experiment 3.

The critical new data are the quality of memory judgments—"remember" vs "know" response. A two-way ANOVA of the "remember" responses showed a significant interaction, $F(1,23) = 12.55$, $MSE = .007$, $p < .005$, $\eta^2_p = 0.353$. There was a significant main effect of item depth of processing, $F(1, 23) = 10.27$, $MSE = .004$, $p < .005$, $\eta^2_p = .309$, but no effect of response type, $F(1,23) = 1.65$, $MSE = .053$, $p > .20$, $\eta^2_p = .067$. Subsequent tests demonstrated significantly more "remember" responses to words from the test of deeply imaged items as compared to shallowly imaged items, $F(1, 23) = 31.36$, $MSE = .004$, $p < .001$, $\eta^2_p = 0.577$. In sharp contrast, there was no significant levels-of-processing effect for items given "know" responses, $F < 1$. These remember/know data are shown in Table 4.1. When the independent remember/know procedure (Yonelinas, 2002) was applied to the "know" responses, the contribution of these responses was higher overall (.35 and .34 for "deep" and "shallow" foils, respectively), but did not differ across foil types.

Therefore, Experiment 5 was able to build upon the previous findings that the memory-for-foils effect is robust and generalizable. In addition, this experiment confirmed that deep foils make up a higher proportion of "remember" responses than do shallow foils, consistent with Marsh et al. (2009b). This suggests that an increase in detail is associated with the foils from the test of deeply encoded words relative to those from the test of shallowly encoded words. This is further supported by research by Gallo, Meadow, Johnson, and Foster (2008), who demonstrated that typical levels-of-processing effects are based on recollective distinctiveness from the extra details that are available for items due to deep encoding. The present argument is that such detail is related to imagery of the items as a consequence of reentry into the picture imagery encoding mode.

Table 4.1. Proportions of hits assigned “Remember” and “Know” responses following imagery-based processing

Response Type	Deep Foils	Shallow Foils
“remember”	.27	.17
“know”	.27	.29

4.2 Experiment 6a: Lexical Decision

The results of Experiment 5 provide some convincing support for the source-constrained retrieval explanation of the memory-for-foils effect. But one might still ask whether foils on a test of deeply encoded items show an overall efficiency of processing that is simply a byproduct of more efficient processing of their deep counterparts. That is, could something about deep processing other than the creation of a mode that can be re-enacted be driving the memory-for-foils effect? To ensure that some sort of general advantage is not the source of the benefit, Experiment 6a substituted a lexical decision task for the deep/shallow processing tasks in the final phase.

If foils from the test of deeply encoded items have a general processing advantage relative to those from the test of shallowly encoded items, then this should be apparent on almost any memory measure. In the case of lexical decision, then, word decisions should be faster for the deep foils as compared to the shallow foils. However, if deep foils were processed under the same encoding mode as their target counterparts, and it is re-entry into this mode that underlies the effect, then there should be no such benefit on the lexical decision task because it is unrelated to that deep encoding.

4.2.1 Method

Participants. A total of 26 undergraduate students from the University of Waterloo (17 female, 9 male) participated for credit. Their mean age was 20.1 years (SD=1.67). Four of the participants were removed from all analyses due to performing in the final phase more than two standard deviations slower than the mean response time for that phase.

Materials. The stimulus words were identical to those used in Experiment 1. Nonwords were compiled using the ARC nonword database (www.maccs.mq.edu.au/~nwdb/nwdb.html; Rastle, Harrington, & Coltheart, 2002). Nonwords were 4–8 letters long and matched with the words on letter length frequency.

Procedure. Participants were tested individually and completed the entire experiment in approximately 30min. Words were randomly assigned to four new lists of 36 words each for each participant. Similarly, nonwords were randomly assigned to two lists of the same size. In addition, each task began and ended with an additional three words (or nonwords) to minimize primacy and recency effects; these buffer items were not included in any analyses. The order of the tasks within each of the phases was counterbalanced across participants.

In the study phase, participants performed deep- and shallow-encoding tasks on separate word lists, identical to the procedure used in Experiment 1.

In the recognition phase (Test 1), participants performed the test precisely as in Experiments 1.

In the judgment phase (Test 2), participants performed a lexical decision task (i.e., "Is the item a word?"). The items on this test consisted of half of the foil items from each of the recognition test lists (18 from the test of deeply encoded items and 18 from the test of shallowly encoded items), intermingled with an equal number of nonwords (72 items in total). To parallel as closely as possible the procedure of Experiment 2, this task was repeated in exactly the same way, with the remaining words from the deep and shallow test lists and a new set of nonwords. Participants were instructed to respond as quickly and as accurately as they could by pressing 1 or 0 on the keyboard. Because there were no methodological differences between the two lexical decision blocks, the data were combined prior to analysis.

4.2.2 Results and Discussion

Recognition test. As before, participants performed well on the recognition test of the initially studied lists (overall hits = .76) and readily discriminated these studied words from new words (overall false alarms = .13). These results are displayed in Figure 4.2a. A paired-samples *t*-test showed that participants had better overall memory for deeply encoded as compared to shallowly encoded words, $t(21) = 10.66$, $p < .001$. This was true for hits, $t(21) = 12.60$, $p < .001$, and showed a mirror effect for false alarms—a greater number of false alarms for shallowly than for deeply encoded words, $t(21) = 2.71$, $p < .01$. Therefore, participants were effectively using the two encoding techniques, resulting in the typical levels-of-processing effect reported by Jacoby et al.

(2005a) and by Marsh et al. (2009b), and seen in the previous experiments in this thesis.

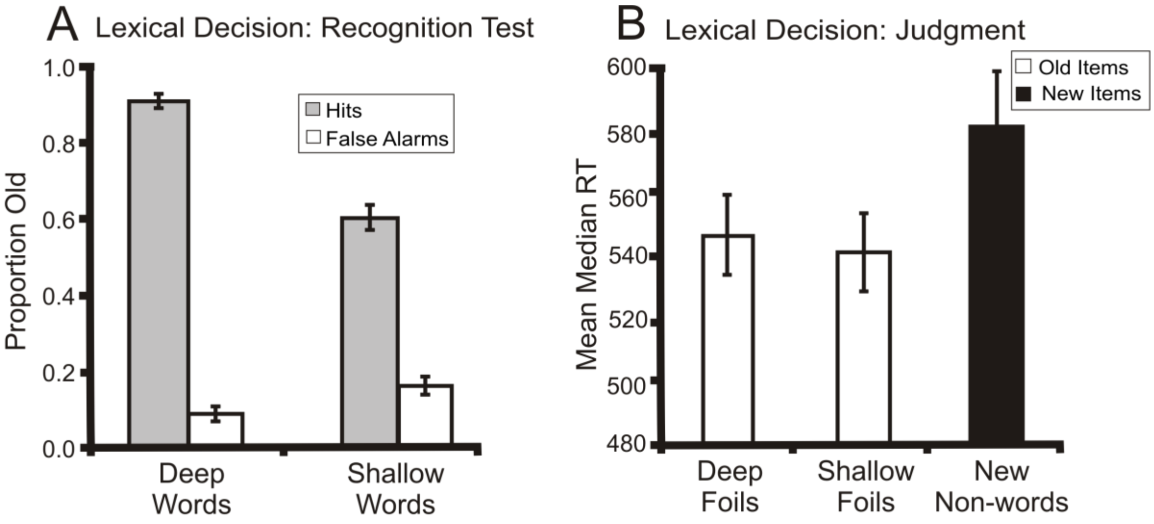

Figure 4.2. Experiment 6a: Manipulating type of decision at test. (a) Recognition data from Test 1, demonstrating enhanced memory following deep encoding as compared to shallow encoding. (b) Reaction time data for Test1 foils on the lexical decision task, demonstrating no significant difference for "deep" and "shallow" foils. Error bars represent the standard errors of the corresponding means.

Lexical decision judgment task. Following one-way ANOVAs, planned contrasts were conducted on the means of the participant median response times for the lexical decision task, which formed the final phase of the experiment. Two contrasts were performed, the first examining priming for the previously seen foils and the second examining whether priming differed between the two types of previously seen foils. There was a significant main effect across the three conditions—deep foils, shallow foils, and new nonwords, $F(2, 42) = 11.80$, $MSE = 885.8$, $p < .001$, $\eta^2_p = .360$. The first contrast demonstrated the routine finding that old words were responded to more quickly than new nonwords, $F(1, 21) = 12.77$, $MSE = 9,672.6$, $p < .01$, $\eta^2_p = .378$. As expected, the second planned contrast resulted in no difference between foils from the test of deeply encoded words and foils from the test of shallowly encoded words, $F(1, 21) = 2.00$, $MSE = 318.9$, $p = .17$, $\eta^2_p = .087$; see Figure 4.2b. Indeed, the observed difference was in the wrong direction, with respect to the hypothesis that deep items should always outperform shallow items. Therefore, foils first encountered among deeply encoded words did not incur any benefit over those first encountered among shallowly encoded items on a measure unrelated to the retrieval mode under which they were thought to be processed.

In sum, words that had been experienced as foils among target words that had been deeply processed at study were not responded to any faster on a subsequent judgment task that did not require the same deep processing. Evidently, if the processing is not the same as during the initial encounter with a word and on the final judgment task, deeply encoded items do not accrue a benefit. Therefore, there is no evidence for some form of general benefit for items processed within a deep context. Based on the constrained retrieval idea and the concept of response mode, the hypothesis offered here is that the benefit for the deep foils should be specific to the judgment that was during the study phase, entirely consistent with those foil items having been processed in the same way as their target counterparts. This was tested in the final experiment of the thesis.

4.3 Experiment 6b: Speeded Judgment

The final experiment was designed with the goal of providing a more direct index of processing mode reinstatement at the time of test. The reasoning was that having prior experience at processing an item in a particular way (or in a particular context) should promote faster processing of that same item within that same context as compared to within a different context. Such encoding specificity should, then, be evident only on a test that is highly related to the original mode of study.

To test this idea, Experiment 6b returned to the typical levels-of-processing study manipulation (i.e., pleasant/unpleasant and “a”/no “a” decisions), for optimal connection to the previous literature. But critically, the final test was changed. In place of the usual recognition test of former foils—Test 2—another speeded judgment test was substituted. This time, the judgment involved repetition of the initial encoding question from the study phase, but carried out now on the foil items from Test 1. Half of the foils from the test of deep items and half of the foils from the test of shallow items were presented together with new items for a pleasantness judgment; the same was done for the letter “a” judgment. The prediction was that if the foils that had accompanied deep targets had been processed deeply (i.e., for pleasantness), whereas the foils that had accompanied shallow targets had not been processed deeply, then only the deep foils would be faster for participants to judge on the pleasantness judgment task, because only they had effectively already been processed deeply in terms of their pleasantness. Participants were not informed that some test item on this pleasantness judgment task would be old and some would be new, so effectively this was an indirect test, unlike the direct recognition test previously used.

If the memory-for-foils effect were a consequence simply of the former foils having been associated with deeply encoded items, it is unlikely that those items would be faster on a subsequent speeded performance test involving the original deep-encoding question. If, however, the deep foil items underwent processing within the same context as their old counterparts during Test 1, then these former foils should be faster to process with respect to pleasantness (the basis of the original deep judgment) than should the shallowly encoded items.

I did not expect a complementary benefit on the shallow judgment task favoring foils that had accompanied shallowly encoded items on the first test because of their relatively weak encoding, and also because I suspected that shallow encoding would not have been sufficient to produce a unique encoding mode that could be successfully reinstated. Nevertheless, to test the alternative hypothesis that accompanying deeply processed items on a prior test always leads to improved memory for foils, I did examine this context by having half of the deep and shallow foils appear on a “contains the letter a” judgment task.

4.3.1 Method

Participants. A total of 41 undergraduate students from the University of Waterloo (24 female, 17 male) participated for credit or remuneration ($5). The mean age was 20.8 years (SD = 3.3). The data of 3 participants were discarded from all analyses due to performing more than two standard deviations slower in the final phase than the mean response time performance for that phase.

Materials. The stimulus words were identical to those used in Experiment 1.

Procedure. Participants were tested individually and completed the entire experiment in approximately 30 min. Words were randomly assigned to six new lists of 36 words each for each participant. In addition, each task began and ended with an additional three words to minimize primacy and recency effects; these words were not included in any analyses. The order of the tasks within each of the phases was counterbalanced across participants.

The study and recognition phases were identical to Experiment 1.

In the judgment phase, there were two subtasks: pleasantness judgment and letter “a” judgment, which were counterbalanced across participants. For pleasantness judgment, participants repeated the original deep-encoding question used at study (“Is the item pleasant or unpleasant?”) for half of the foil items from each of the recognition test lists (18 from the test of deeply encoded items and 18 from the test of shallowly encoded items) intermingled with 36 new items (72 words in total). The remaining deep and shallow foil items from the first recognition phase were mixed with another set of new items, and for these participants responded to the same shallow-encoding question used during study (“Does the word contain an ‘a’ or no ‘a’?”). Thus, both “deep” and “shallow” foils were tested with each judgment task such that half of the foils were aligned in terms of the context of their recognition judgment, and half were not. Additional instructions requested that participants respond as quickly as possible while performing as accurately as they could. As before, they responded by pressing 1 or 0 on the keyboard. Participants were never instructed as to the nature of the words; that is, they were never told that old words would be appearing among the items during these decision tasks.

4.3.2 Results and Discussion

Recognition test. Participants performed well on the recognition test of the initially studied lists (overall hits = .75) and readily discriminated these studied words from new words (overall false alarms = .21). These results are displayed in Figure 4.3a. A paired-samples *t*-test showed that participants had better overall memory for deeply encoded as compared to shallowly encoded words, $t(37) = 10.91$, $p < .001$. This was true for hits, $t(37) = 8.86$, $p < .001$, and showed a mirror effect for false alarms—a greater number of false alarms for shallowly than for deeply encoded words, $t(37) = 4.31$, $p < .001$. Therefore, participants were effectively using the two encoding techniques, resulting in the typical levels-of-processing effect reported by Jacoby et al. (2005a) and by Marsh et al. (2009b), and also replicated Experiment 1 in this thesis.

Judgment task. Following one-way ANOVAs, planned contrasts were conducted on the means of the participant median response times for each of the judgment tasks, which together formed the final phase of the experiment. For each judgment task, there were two contrasts, the first examining priming for the previously seen foils, and the second examining whether priming differed between the two types of previously seen foils.

Shallow judgment task. On the shallow judgment task, the three conditions—deep foils, shallow foils, and new words—did not differ from each other, $F < 1$. Not surprisingly, therefore, neither planned comparison was significant, both *F*s < 1 (for shallow vs deep; Figure 4.3b). Therefore, priming did not occur either overall, for old versus new words, or differentially, for shallow versus deep test foils. I suspect that the processing carried out in judging whether words contain the letter "a" is so limited that participants cannot benefit from reinstating the vowel-based shallow mode, if indeed there actually is such a mode. This condition was included just for completeness, as I did not expect any differential priming of items from the different test lists.

Deep judgment task The task of principal interest was the deep judgment task, since the findings of Jacoby et al. (2005a) and Marsh et al. (2009b) had suggested that this mode of processing can be reinstated. If the foils presented among deep targets on the recognition test were processed like the deep targets had been during study (i.e., for pleasantness), this should result in more priming of that same judgment for the deep foils relative to the shallow foils. There was a significant main effect across the three conditions—deep foils, shallow foils, and new words, $F(2, 74) = 7.46$, $MSE = 1{,}666.4$, $p < .001$, $\eta^2_p = .167$. The first contrast showed an overall priming effect: Old words were

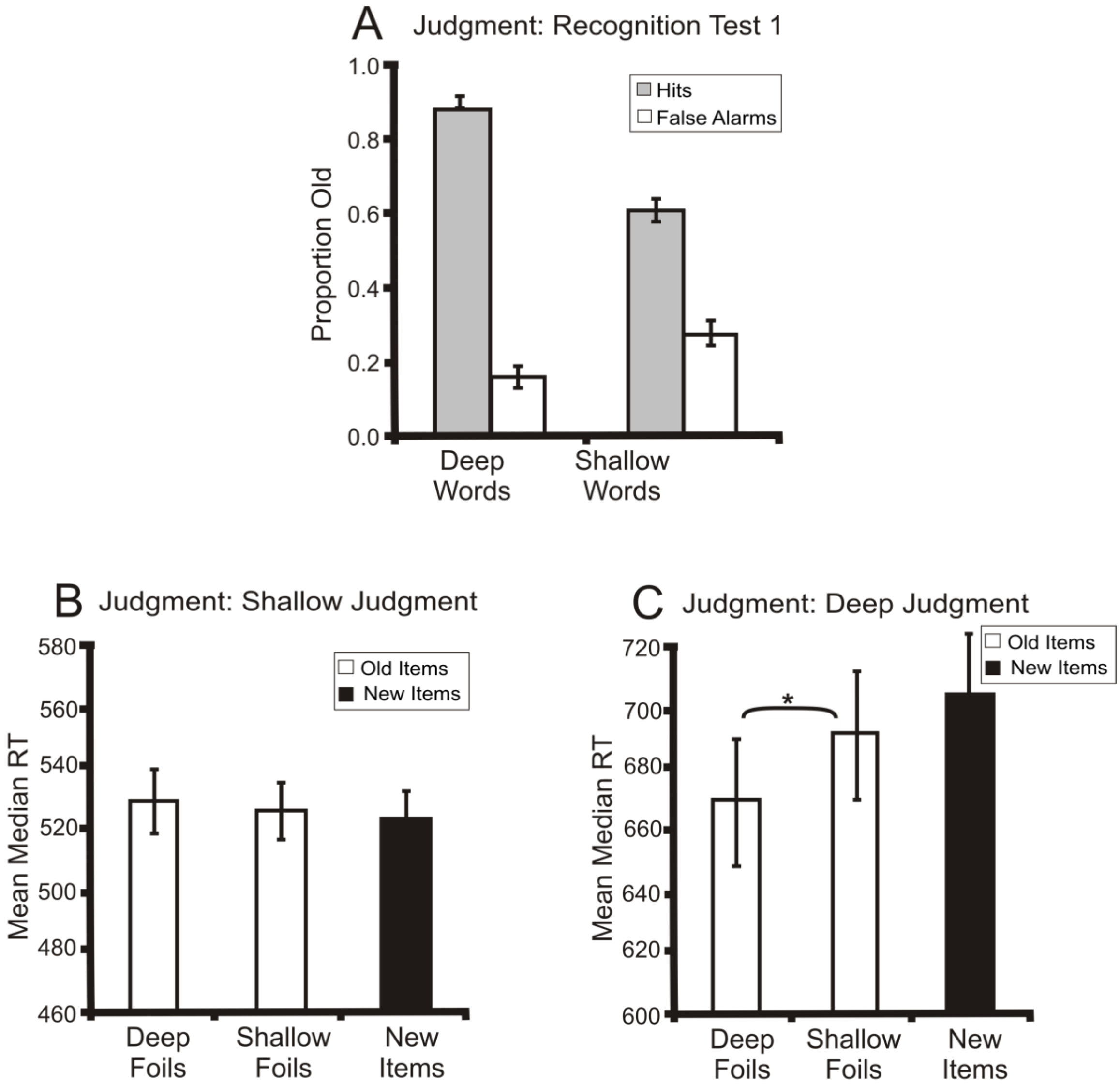

Figure 4.3. Experiment 6b: Manipulating type of decision at test. (a) Recognition data from Test 1, demonstrating enhanced memory following deep encoding as compared to shallow encoding. (b) Reaction time data for Test 1 foils on the shallow decision task demonstrating no significant difference for "deep" and "shallow" foils. (c) Reaction time data for Test 1 foils on the deep decision task, demonstrating a "deep" foil advantage over "shallow" foils. Error bars represent the standard errors of the corresponding means.

responded to more quickly than new words, $F(1, 37) = 13.0$, $MSE = 7{,}410.9$, $p < .001$, $\eta^2_p = .260$. The second planned contrast was the crucial test and did indeed demonstrate that participants were faster at making the pleasantness judgment for the foils from the test of deeply encoded words than for the foils from the test of shallowly encoded words, $F(1, 37) = 4.11$, $MSE = 4{,}195.2$, $p < .05$, $\eta^2_p = .100$; see Figure 4.3c.

In sum, words that had been experienced as foils among target words that had been deeply processed at study benefited on a subsequent judgment task that required the same deep processing. This was not simply general priming from prior experience, because words experienced as foils among target words that had been shallowly processed at study showed reliably less priming. Contrasting Experiments 6a and 6b demonstrates that the benefit for the deep foils was specific, consistent with these items having been processed in the same way as their target counterparts. This provides direct evidence in support of the idea of source-constrained retrieval because, for such a benefit to occur, the words would have to have been associated with that relevant type of processing in a prior encounter—through re-entry into the encoding context during the prior recognition test.

CHAPTER 5: GENERAL DISCUSSION

Encoding and retrieval processes are often conceptualized as being separate entities. However, transfer appropriate processing and, more recently, the memory-for-foils phenomenon, would suggest otherwise. Indeed, Hintzman (2011) dubs this “process-pure” assumption with respect to encoding/retrieval one of the prevalent fallacies in the study of memory. The main tenet of this thesis is that these processes are, instead, part of one overarching process. While encoding may be about the laying down of new memories and retrieval about the recovery of old ones, it is inescapable that the encoding and retrieval of memories are intimately linked. For them to be linked, any encoding is done within the context of prior memories and is therefore another retrieval episode; any retrieval done within the context of new experiences is therefore another encoding episode. The purpose of this thesis is to demonstrate one way in which this occurs.

5.1 Summary of research findings

Jacoby et al. (2005a; see also Jacoby et al., 2005b; Marsh et al., 2009b; Shimizu & Jacoby, 2005) demonstrated that the way in which targets are processed on a recognition test can influence subsequent memory for the accompanying distractors. Specifically, distractor words that had appeared among target words that had been semantically encoded during an initial study phase were subsequently better recognized than distractor words that had appeared among targets that had been encoded non-semantically during initial study. Experiment 1 of this thesis reports a faithful replication of this basic finding. We also know from the work of Marsh et al. (2009b) that this effect is not simply the consequence of differential strength of encoding, as they showed by manipulating number of presentations in their Experiment 3. Experiment 2 reports a faithful replication of this finding as well. Experiment 3 generalizes this memory-for-foils effect from the levels-of-processing manipulation used previously to a novel imagery encoding manipulation. Words that people did not intend to learn nevertheless benefited on a later memory test when they were experienced among other, previously elaborated, words; we now know that this occurs using two of the most widely studied modes of elaboration: levels-of-processing and imagery.

Whereas the levels-of-processing mode is based on the degree of semantic—as opposed to perceptual—processing, the elaboration brought about by imagery as an encoding mode certainly appears to have a different basis. Imagery is not equivalent to semantic processing, invoking as it does perceptual elements of what is imaged (see Paivio, 1971, 1995, 2007). But imagery is a coherent mode of processing in the same sense that deep semantic processing is: Both are readily engaged—and re-engaged—ways of thinking about what is presented. This is why I reasoned that deep versus shallow imagery should also be capable of inducing and re-inducing differential modes of processing.

When this same study was extended, in Experiment 5, to include a judgment regarding the quality of the memories for the foils in the final recognition test, the results demonstrated that more detail-based memories accompanied “deep” foils than “shallow” foils during that test. This study was a stepping stone to providing direct evidence for the interaction between encoding and retrieval processes. Indeed, while participants were undergoing retrieval processing for old, deep items in the first test, new items were concurrently being encoded with more details than would be expected in the absence of a deep context associated during retrieval. Such detail-rich memories suggest a link between the depth of encoding of the old items and the way in which accompanying new items are encoded during the test.

Experiment 6 fits the key piece to the puzzle, and as such is the centerpiece of the thesis. Here, the question of whether it would be possible to obtain more direct behavioral evidence of re-entering the original encoding mode was addressed. If, during Test 1, the foils are re-processed with respect to the original mode of processing of the accompanying targets, then that should be evident when the foils subsequently must be processed in terms of that original mode. To test this hypothesis, Experiment 6b ended not with a recognition test but with the same judgment task that was used during study. By showing that people were faster to respond on a pleasantness judgment task to foils from the test of deeply encoded words, Experiment 6b demonstrated that these words were encoded within that same deep context. Further, we know from Experiment 6a that this benefit for "deep" foils is not due to a general processing benefit for items associated with deeply encoded words but, instead, only occurs within the context of the original encoding task. Therefore, during retrieval, participants do appear to re-enact the encoding task.

What seems to be essential to benefit memory for the foils is that encoding involves differentiable modes of processing being applied to the two sets of words during study, and that the reinstatement of these same modes of processing—separately—be accomplished at the time of the first recognition test. If both conditions are met, and if encoding was initially done more elaborately, then the foils also receive more elaborative encoding—the same more elaborative encoding—and are better remembered. In the framing of Jacoby et al. (2005a), the beneficial encoding mode is reinstated on the first recognition test, in accordance with the transfer appropriate processing principle (Morris et al., 1977; Roediger, 1990). This is what Jacoby et al. referred to as "source-constrained retrieval."

It appears, then, that we unintentionally process items on a second occasion in much the same way as we processed them on the first occasion, even without any explicit requirement to do so. This is not surprising: It is in accord with the idea of transfer appropriate processing (Morris et al., 1977), which shows that retrieval is best when the processes engaged during retrieval match those that were engaged during encoding. This phenomenon certainly meshes well with the proceduralist view of memory (Kolers, 1973; Kolers & Roediger, 1984) where memory is conceived more as a byproduct of processing during encoding and retrieval. In this view, it is actually the *process* of how things are put into and taken out of 'memory' that is most important rather than the *intention* to do so (see also Craik, 1983).

It is worth noting, however, that the benefit of transfer appropriate processing stems from processing during *study*, whereas the benefit of source-constrained retrieval results from reprocessing of items during *test*. Such processing reinstatement optimizes retrieval success when it provides a coherent encoding mode, "spilling over" onto other items processed contiguously, even without any intention to learn them. This highlights that there is indeed a mode of processing that is active across trials during retrieval.

This agrees with the proceduralist analysis that there is very substantial overlap of the processes involved in encoding and retrieval. Instead of thinking of retrieval as separate and distinct from encoding, retrieval could more parsimoniously be regarded as another encoding event.

Indeed, in Tulving and Thompson's (1973) concept of "synergistic ecphory" as part of the encoding specificity principle, the type of retrieval cues that are specific to the encoded items are fundamental to the success of the memory retrieval. That is, there is more information available in memory at any given time than is accessible. Analogously to how forgetting can be caused by the lack of availability of appropriate cues, encoding can also benefit from the appropriate type of retrieval strategy.

Even when more complex stimuli are used, as was the case in Experiment 4 using phrases, the memory-for-foils effect is still obtained. The amount of information contained within one sentence is compounded relative to words. It is rare that we have to remember words in isolation. Therefore, phrases are, therefore, complex stimuli that have increased applicability when considering the broader context of memory within our daily lives. Thus, encoding and retrieval processes do, in all likelihood, work together to aid in creating better, more highly associated memories outside of the context of the laboratory setting. Interestingly, the sentences in Experiment 4 were unrelated—the only link is that of the study manipulation (i.e., "How believable is this statement?"). Thus, even with a multitude of information, the way that it is stored (and therefore retrieved) has great consequences for new, incoming information.

This thesis has demonstrated a link between encoding and retrieval processes: The way that old items are retrieved has a direct and measurable influence on the success of encoding of new items. The experiments have also shown that this influence is not restricted to a single mode of processing. Re-invoking the encoding processes (or modes) during retrieval permits all items on the recognition test (including the new items) to undergo that processing, with the same benefits to memory for the new items as were observed for the originally studied items. The new items are thus encoded using a retrieval process that increases the likelihood of richer encoding, and produces measurable facilitation in speed of subsequent processing. Importantly, these results demonstrate that the mode of processing engaged during encoding, and reinstated during retrieval, has substantial effects on the encoding of new information, thereby helping to specify how encoding and retrieval are linked.

This link is also starting to emerge in neuroimaging studies. While no neuroimaging results are presented here, my findings can be tied with recent evidence that some of the regions that are activated during encoding are re-activated again at retrieval (Danker & Anderson, 2010; Nyberg, Habib, McIntosh, & Tulving, 2000; Nyberg, Petersson, Nilsson, Sandblom, Aberg, & Ingvar, 2001; Wheeler & Buckner, 2003; Woodruff, Johnson, Uncapher, & Rugg, 2005). Nyberg and colleagues (2000) associated words with sounds during encoding using PET. At retrieval, although only the words were presented, auditory brain regions were also active, suggesting that all aspects of the encoding experience were retrieved. Using fMRI, Skinner, Grady, & Fernandes (2010), also found reactivation during retrieval, this time in face-related regions (i.e., the fusiform gyrus), for words that had been studied along side faces as compared to those that had not. Although this area needs further investigation, these studies support a strong link between encoding and retrieval processes.

5.2 Potential Criticisms

One might argue that these conclusions all derive from the context of a relatively specific set of parameters and a single paradigm, and consequently might not hold in a different context. However, changing the type of stimuli (i.e., words to phrases) and the type of study manipulation (i.e., semantic elaboration to imagery) provides support for the phenomenon's generalizability. But the question still might arise as to the role that this effect plays in everyday life. The argument in this thesis is that it is likely that there is a strong role for this type of interaction between encoding and retrieval in common memory problems. In addition to my earlier example of remembering a grocery list, we often use the context of encoding during attempts at retrieval, even for something as banal as remembering an actor's name: We first may go through the movies in which they have appeared, in an effort to spur the memory for their name.

With unintentional memories, encoding and retrieval are still likely to co-operate such that incoming information is immediately linked to pre-existing information in memory. The “butcher on the bus” phenomenon provides a salient example of this. This phenomenon (often credited to Mandler, 1980, but perhaps first broached by Osgood, 1953, p. 550), shows how encountering someone in an unexpected context results in difficulty recognizing that person, who otherwise would be easily recognized. Unintentionally encoded information (e.g., the butcher shop) is truly necessary to retrieve the memory of who the person is. This shows how important the context of encoding is to retrieval; retrieval is seriously impeded if its context does not align with the encoding context.

Another potential criticism is the lack of finding of a shallow processing advantage in experiment 6b. When words that had been in the shallow recognition test and words that had been in the deep recognition test both appeared on a shallow judgment task—the same one used during initial study—the words from the shallow recognition test were not processed faster than those from the deep test. There are several potential explanations for this. One is that it is possible that the effect was smaller and harder to obtain with that task because the differences between items is much smaller (all drawn from a finite set of letters). Another possible reason is that when the encoding mode is not strong enough, retrieval plays a much smaller role during encoding; for items within a weak context, there may be an associated broader context (i.e., the lab session) but not something more specific (i.e., shallow processing). My preferred explanation is simply that judging whether a word contains a particular letter is not an analysis likely to be replayed during subsequent testing—this task does not constitute a retrieval mode that can be re-enacted.

A further criticism might pertain to the likelihood of the phenomenon of retrieval mode occurring within the context of everyday memories. I discuss this idea further below but, essentially, I argue that the more relevant question would be when it would *not* occur in the context of our everyday memories.

5.3 Future Directions

One remaining question pertains to the expansion upon the question of applicability of the memory-for-foils effect. Are people likely to use it as a matter of course in their everyday lives? Is the context of retrieved items automatically associated with novel ones? For example, on multiple choice tasks used by so many university classes, are students actually integrating the novel (i.e., false) information when they have studied for the test using deep encoding techniques? If so, it seems that we are actually causing harm as well as good to the best students by increasing the amount of false information being incorporated into their memories of the course material. This would suggest that we should move away from this type of testing. Indeed, there is some evidence in the literature on the testing effect—the benefits of testing—for exactly this sort of cost (see, e.g., Butler & Roediger, 2008; Roedgier & Marsh, 2005; Marsh, Agarwal, & Roediger, 2009a).

An important piece of the puzzle that still needs further investigation is that of neuroimaging evidence of an overlap between encoding and retrieval processes. As noted earlier, it is within this realm that the evidence for the separability of these processes has been emphasized. While there is some emerging evidence for such an overlap (Danker & Anderson, 2010; Nyberg et al., 2000; Nyberg et al., 2001; Wheeler & Buckner, 2003; Woodruff et al., 2005), I believe that, in the context of neuroimaging, the memory-for-foils paradigm could be a very suitable tool for addressing this question. If, for example, we were to find the same areas to be activated during the original encoding session as for the foils within the first test, it would be convincing evidence for a role for both processes occurring during the test, especially because the study manipulation and the test place very different requirements on the participants. In particular, activation would be expected in areas involved in visual imagery (i.e., the prefrontal cortex, parahippocampal cortex and occipital regions; Johnson & Rugg, 2007). In addition to the use of the memory-for-foils paradigm, it is important to use newer analyses that allow us not only to investigate singular areas of activation, but also to discover areas that work together within the network, thereby providing a possible missing link. Prior research has faced difficulty in designing a study that would offer such a clear overlap between retrieval and encoding, which has made drawing any conclusions about their interactive nature more difficult.

Another question that remains is whether the depth of encoding needs to be associated with the study items directly or whether that depth—that mode—could somehow be induced during the test, if that would provide the same benefits as seen with the memory-for-foils paradigm. For example, if all study items could be categorized in a way that would not be obvious during study, if that categorization was then provided prior to the test, would the new foils still be linked? Other contextual dimensions such as physical location and mood for the separate lists could also speak to the issue of what type of context plays a role.

As with many other memory phenomena (for example, see McConkie & Currie, 1996; Simons & Chabris, 1999), attention is likely a necessary component. We know that the relation between encoding and retrieval on a levels-of-processing task can be interrupted when attention is divided, especially during encoding (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Naveh-Benjamin, Craik, Gavrilescu, & Anderson, 2000). Division of attention during the memory-for-foils paradigm has yet to be tested, but this would make an interesting next step, particularly were attention to be divided during the first recognition test, potentially preventing the implementation of a retrieval mode during the test, and thereby reducing or eliminating the memory-for-foils effect.

Within the studies reported here, I have demonstrated a powerful—and very specific—influence of retrieval on encoding processing. An important next step would be to expand upon the converse effect—the effect of encoding on retrieval. In addition to transfer-appropriate processing, there is some evidence that this influence does occurs, given that the *way* in which we encode things has a powerful effect on how well items are later retrieved (e.g., levels of processing, Craik & Lockhart, 1972; and remember/know; Gardiner, 1988; Rajaram, 1993). Moreover, the generation effect—where items that are generated from clues are better remembered than those that are simply read (Slamecka & Graf, 1978; Bertsch, Pesta, Wiscott, & McDaniel, 2007)—suggests a strong role for retrieval in successful encoding. It also could likely be shown that retrieval is impaired based on a conflicting encoding manipulation, supporting the notion that the two processes work together.

5.4 Conclusions

Within the context of the memory-for-foils paradigm, the empirical work in this thesis has demonstrated a very intimate link between encoding and retrieval processes. The existence of this link argues for a fundamental shift in the way that the processes are conceived and studied, as Hintzman (2011) has recently convincingly contended, writing that there are “two implicit assumptions that are false: that retrieval does not occur during the study phase, and that encoding does not occur during the test phase” (p. 256). Creating artificial divisions between encoding and retrieval leads to less than ideal perceptions of the processes. To truly understand memory as a whole, the way in which these sub-processes function together is, I believe, the more fundamental question. As part of one larger process, encoding and retrieval work together to build an efficient and effective memory system—one in which memories are well organized and, thus, easily accessible. We know that memories are never perfect replicas of past events, but rather, are reconstructed or recreated (Bartlett, 1932). Therefore, it is part of the fundamental process that new memories are integrated well not only with the appropriate external context, but also into our existing personal memories. This integration acts to strengthen the memories and increase their value because they can provide some extra context and meaning. Without this, nothing would be linked and learning would be very difficult indeed.

In considering the usefulness of encoding/retrieval interactions, there appears to be a very powerful and useful mnemonic that is not commonly suggested but that could perhaps benefit people were they encouraged to use it. If people are instructed directly to try to retrieve each item by repeating the study manipulation on them, this would benefit not only memory for those items, but additionally memory for the co-occurring novel items. This is only one potential positive outcome for this link between encoding and retrieval but, regardless, this link could have many important ramifications with respect to the way that we conceptualize memory and study it.

Memory processes hold a great deal of fascination for many researchers with the goal of understanding human cognition, myself included. But much of society gives these processes little thought until they fail to function. There is so much complexity when thinking about how we encode and retrieve memories within behavioural studies of cognition alone, without even brushing the edges of the neuroscientific realm. Every new piece of the puzzle continues to bring with it more questions. Despite the common colloquialisms about memory that refer to encoding and retrieval as separate, they clearly are not. It is the personal integration of events across our present and our past that makes the processes that underlie our understanding of the world truly fascinating—and extremely elusive. The principal argument in this thesis is that we can only understand our own memory processes if we view them as interdependent, and stop treating them as isolated from each other. To learn is to remember; to remember is to learn.

REFERENCES

Aggleton, J.P., McMackin, D., Carpenter, K., Hornak, J., Kapur, N., Halpin, S., Wiles, C.M., Kamel, H., Brennan, P., Carton, S., & Gaffan, D. (2000). Differential cognitive effects of colloid cysts in the third ventricle that spare or compromise the fornix. *Brain, 123*, 800–815.

Bai, F., Zhang, Z., Watson, D.R., Yu, H., Shi, Y., Yuan, Y., Zang, Y., Zhu, C., Qian, Y. (2009). Abnormal functional connectivity of hippocampus during episodic memory retrieval processing network in amnestic mild cognitive impairment. *Biological Psychiatry, 65,* 951-8.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, UK: Cambridge University Press.

Bentin, S., & Moscovitch, M. (1988). The time course of repetition effects for words and unfamiliar faces. *Journal of Experimental Psychology: General, 117*, 148-160.

Bertsch, S., Pesta, B., Wiscott, R., & McDaniel, M. (2007). The generation effect: A meta-analytic review. *Memory & Cognition, 35*, 201-210.

Butler, A.C., & Roediger, H.L., III (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36,* 604-616.

Cansino, S., Trejo-Morales, P., & Hernández-Ramos, E. (2010). Age-related changes in neural activity during source memory encoding in young, middle-aged and elderly adults. *Neuropsychologia, 48*, 2547-2549.

Cipolotti, L., & Bird, C.M. (2006). Amnesia and the hippocampus. *Current Opinion in Neurology, 19*, 593-8.

Craik, F.I.M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society B, 302*, 341–359.

Craik, F.I.M., Govoni, R., Naveh-Benjamin, M., & Anderson, N.D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125*, 159–180.

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

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28*, 923–938.

Danker, J.F., & Anderson, J.R. (2010). The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin, 136,* 87-102.

Dannhauser, T.M., Shergill, S.S., Stevens, T., Lee, L., Seal, M., Walker, R.W., & Walker, Z. (2008). An fMRI study of verbal episodic memory encoding in amnestic mild cognitive impairment. *Cortex*, *44*, 869-80.

Desgranges, B., Baron, J.C., & Eustache, F., (1998). The functional neuroanatomy of episodic memory: the role of the frontal lobes, the hippocampal formation, and other areas. *Neuroimage, 8,*198-213.

Dewhurst, S.A., & Brandt, K.R. (2007). Reinstating elaborative encoding operations at test enhances episodic remembering. *Quarterly Journal of Experimental Psychology, 60*, 543-550.

Düzel, E., Cabeza, R., Picton, T.W., Yonelinas, A.P., Scheich, H., Heinze, H.J., & Tulving, E. (1999). Task-related and item-related brain processes of memory retrieval. *Proceedings of the National Academy of Sciences, USA, 96,* 1794-9.

Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences ofthe USA*, *94*, 5973–5978.

Ebbinghaus, H. (1885/1962). *Memory: A contribution to experimental psychology.* New York: Dover.

Gabrieli, J.D.E., Brewer, J.B., Desmond, J.E., & Glover, G.H. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science, 276*, 264-266.

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

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language, 58*, 1095-1111.

Gopie, N., & MacLeod, C. M. (2009). Destination memory: Stop me if I've told you this before. *Psychological Science, 20,* 1492-1499.

Hintzman, D. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 47-91). New York: Academic Press.

Hintzman, D.L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth." *Perspectives on Psychological Science*, *6*, 253-21.

Hourihan, K. L. (2008). *The power of optimal encoding: Distinctiveness and differentiation defeat directed forgetting.* Dissertation submitted to the Department of Psychology, University of Waterloo.

Hourihan, K. L., & MacLeod, C. M. (2011). *The power of optimal encoding: Distinctiveness effects in directed forgetting.* Manuscript submitted for publication.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory, 14*, 47-53.

Hupbach, A., Gomez, R., & Nadel, L. (2009): Episodic memory reconsolidation: Updating or source confusion? *Memory*, *17*, 502-510.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005a). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, *12*, 852-857.

Jacoby, L. L., Shimizu, Y., Velanova, K., & Rhodes, M. G. (2005b). Age differences in depth of retrieval: Memory for foils. *Journal of Memory and Language*, *52*, 493-504.

James, W. (1890). *Talks to teachers on psychology: And to students on some of life's ideals.* New York: Henry Holt and Company.

Java, R.I., Gregg, V.H., & Gardiner, J.M. (1997). What do people actually remember (and know) in “remember/Know” experiments? *European Journal of Cognitive Psychology, 9*, 187-197.

Johnson, J.D., & Rugg, M.D. (2007). Recollection and the reinstatement of encoding-related cortical activity. *Cerebral Cortex, 17*, 2507–2515.

Karpicke, J. D., & Zaromb, F. M. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal of Memory and Language*, *62*, 227-239.

Kolers, P. A. (1973). Remembering operations. *Memory & Cognition, 1*, 347-355.

Kolers, P. A., & Roediger, H. L., III (1984). Procedures of mind. *Journal of Verbal Learning & Verbal Behavior, 23,* 425-449.

Koriat, A. (2000). Control processes in remembering. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 333-346). New York: Oxford University Press.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET ctivations of memory encoding and retrieval: The HIPER model. *Hippocampus, 8*, 313-322.

Macken, W.J. (2002). Environmental context and recognition: The role of recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 153-161.

Mandler, G. (1980). Recognizing: the judgment of previous occurrence. *Psychological Review*, *87*, 252– 271.

Mark, R.E., & Rugg, M.D. (1998). Age effects on brain activity associated with episodic memory retrieval: An electrophysiological study. *Brain, 121***,** 861–873.

Marsh, E.J., Agarwal, P.K., & Roediger, H.L. (2009a). Memorial consequences of answering SAT II questions. *Journal of Experimental Psychology: Applied, 15,* 1-11.

Marsh, R. L., Thadeus Meeks, J., Cook, G. I., Clark-Foos, A., Hicks, J. L., & Brewer, G. A. (2009b). Retrieval constraints on the front end create differences in recollection on a subsequent test. *Journal of Memory and Language*, *61*, 470-479.

McConkie, G. W. & Currie, C.B. (1996), Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance, 22*, 563–581.

Medved, M.I., & Hirst, W. (2006). Islands of memory: Autobiographical remembering in amnestics. *Memory, 14,* 276–88.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533.

Mulligan, N. W., & Lozito, J. P. (2006). An asymmetry between memory encoding and retrieval: Revelation, generation, and transfer-appropriate processing. *Psychological Science, 17*, 7-11.

Murnane, K., & Shiffrin, R. M. (1991). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 855-874.

Naveh-Benjamin, M., Craik, F.I.M., Gavrilescu, D., & Anderson, N.D. (2000). Asymmetry between encoding and retrieval processes: Evidence from divided attention and a calibration analysis. *Memory & Cognition, 28*, 965–976.

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences, USA, 97,* 11120–11124.

Nyberg, L., Petersson, K.M., Nilsson, L.G., Sandblom, J., Aberg, C., & Ingvar, M. (2001). Reactivation of motor brain areas during explicit memory for actions. *NeuroImage, 14,* 521–528.

Nyberg, L., Tulving, E., Habib, R., Nilsson, L.-G., Kapur, S., Houle, S., et al. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *Neuroreport, 7*, 249-252.

Osgood, C.E. (1953). *Method and theory in experimental psychology.* New York: Oxford University Press.

Paivio, A. (1971). *Imagery and verbal processes.* Oxford, UK: Holt, Rinehart & Winston.

Paivio, A. (1995). Imagery and memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 977-986). Cambridge, MA: The MIT Press.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach.* Mahwah, NJ: Lawrence Erlbaum Associates.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, *21*, 89-102.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 non-words: The ARC nonword database. *Quarterly Journal of Experimental Psychology, 55A*, 1339-1362.

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

Roediger, H. L., III (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043–1056.

Roediger, H. L., III, Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels of processing framework. *Memory, 10*, 319-332.

Roediger, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 31, 1155-1159.

Rogers T.B., Kuiper, N.A., & Kirker, W.S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*, 677-688.

Shimizu, Y., & Jacoby, L. L. (2005). Similarity-guided depth of retrieval: Constraining at the front end. *Canadian Journal of Experimental Psychology*, *59*, 17-21.

Simons, D.J., & Chabris, C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28,* 1059–1074.

Skinner, E.I., Grady, S.L., & Fernandes, M.A. (2010). Reactivation of context-specific brain regions during retrieval. *Neuropsychologia, 48*, 156–164.

Slamecka, N.J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 592-604.

Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition 24*, 1379-1396.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College, Columbia University.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352-373.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26,* 1-12.

Tulving, E. (1998). Brain/mind correlates of human memory. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in psychological science, Vol. 2. Biological and cognitive aspects* (pp. 441-460). Hove, UK: Psychology Press/Erlbaum.

Tulving, E., Kapur, S., Craik, F.I., Moscovitch, M., & Houle, S. (1994) Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proceedings of the National Academy of Sciences, U.S.A., 91,* 2016–2020.

Van Damme I, d'Ydewalle G. (2009). Memory loss versus memory distortion: the role of encoding and retrieval deficits in Korsakoff patients' false memories. *Memory, 17,* 349-66.

Ward, J. (2003). Encoding and the frontal lobes: a dissociation between retrograde and anterograde memories. *Cortex, 39,* 791-812.

Wheeler,M. E., & Buckner, R. L. (2003). Functional dissociations among components of remembering: Control, perceived oldness, and content. *Journal of Neuroscience, 23,* 3869–3880.

Woodruff, C. C., Johnson, J. D., Uncapher, M. R., & Rugg, M. D. (2005). Content-specificity of the neural correlates of recollection. *Neuropsychologia, 43,* 1022–1032.

Yonelinas, A P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441-517.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, *12*, 323–339.

APPENDIX A

The most stolen car in Canada is the 2000 Honda Civic.

Spermology is the study or collection of trivia.

Coffee is the world's most popular beverage.

The hottest chili in the world is the habanero.

A cluster of bananas is called a hand.

Bats are the only mammal that can fly.

The average life span of a mosquito is two weeks.

Ants do not sleep.

Amelia Earhart was the second person to fly solo across the Atlantic.

The opposite of a "vacuum" is a "plenum."

Naples is the place of origin of pizza, guitars and mandolins.

Adding milk to tea negates its health effects.

VHS stands for Video Home System.

The Hope Diamond was mined in India.

The Silk Road runs from China to Turkey.

India ink is not from India - it's from China and Egypt.

There are 1,792 steps in the Eiffel Tower.

Cloudy apple juice is healthier than clear juice.

A prairie oyster is actually a calf's testicle.

Racism is a crime in Brazil.

Some roaches can fly.

An octopus has three hearts.

A firefly is not a fly; it's a beetle.

Paper products make up approximately 40 percent of all trash.

Sony started as the maker of rice cookers.

Matthew means "Gift from God."

People of Monaco are called Monegasques.

The average NFL career lasts about 3½ years.

A house mouse feeds 15 to 20 times per day.

Dom Perignon was a Benedictine monk.

Whales have no vocal cords, but they can sing.

The Latin words for apple and evil are the same: malus.

There are more pyramids in Sudan than in Egypt.

A leapling is a person born on February 29th.