Spring 5-15-2015

Using the Encoding Specificity Principle to Assess the Nature of the Secondary Memory Component of Working Memory

Dung Chi Bui
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychology

Dissertation Examination Committee:
Sandra S. Hale, Chair
David A. Balota
Carl Craver
Joel Myerson
Henry L. Roediger, III

Using the Encoding Specificity Principle to Assess the Nature of the Secondary Memory Component of Working Memory

by

Dung Chi Bui

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2015

St. Louis, Missouri

Table of Contents

List of Figures

Acknowledgments

Over the last eight years, I have had the privilege of working with exceptional colleagues on projects that have attempted to answer a wide variety of questions. This dissertation represents my growth and maturity as a researcher, and highlights some of the questions that I have come to believe are important to answer. This work, however, would not have been possible without the remarkable support and guidance that I have received during my time here at Washington University.

I would like to thank my dissertation committee for their time in providing valuable feedback and comments. I would especially like to express my gratitude towards my advisors, Sandy and Joel, for providing endless support, and for taking a chance on me when many others would not. During my time here, they have nurtured me and helped me develop the confidence that I needed to not only grow as a scientist, but as a person. I would also like to thank Roddy and Dave for opening up their time and resources to me, and for providing an instrumental outlet for me to explore new research interests. I chose to come to Washington University because the psychology department was known for encouraging collaboration, and looking back, there is no doubt in my mind that Roddy and Dave are huge reasons why the department has this reputation.

I would not be where I am now if it were not for my mother, who fled from her war-torn home and came to this country with nothing, but made endless sacrifices to give me a chance at receiving a quality education. She has always challenged me to push my boundaries, and believed that I was capable of doing meaningful things with my life, long before I even knew it. Finally, I would like to thank my partner, Laura, who has been amazingly graceful in making sacrifices so that I can pursue my own goals. Her unconditional support for me is more than I deserve, and I could not ask for anyone better than her to be by my side as I embark on my research career.

ABSTRACT OF THE DISSERTATION

Using the Encoding Specificity Principle to Assess the Nature of the Secondary Memory Component of Working Memory

by

Dung Chi Bui

Doctor of Philosophy in Psychology

Washington University in St. Louis, 2015

Professor Sandra Hale, Chairperson

Recent theories and evidence suggest working memory involves secondary memory as well as primary memory. It is unclear, however, if the secondary memory component of working memory is the same as the secondary memory component underlying episodic long-term memory. The present investigation explores this issue by examining whether manipulating encoding and retrieval cues on a short-term memory task produces similar effects as to what is typically seen on episodic long-term memory tasks. More specifically, it is commonly observed on episodic long-term memory tasks that retrieval cues that were not also present during encoding produces worse recall compared to retrieval cues that were present during encoding, as well as worse recall compared to if no cues were presented. Currently, it is unclear whether this finding, known as the encoding specificity principle, would also be observed in short-term memory tasks. In the current investigation, participants engaged in a modified operation span task where they learned weakly related word-pairs ("era : TIME"). During recall, participants were either provided the same cue from earlier in the series ("era"; match cue), a different cue that was not shown earlier in the series but was strongly associated with the target word ("life"; mismatch cue), or were asked to free recall the target word (no cue). Under conditions in which

performance was predicted to rely on secondary memory, performance in the no cue condition was better than the mismatch condition, consistent with the encoding specificity principle (Thomson & Tulving, 1970). Importantly, when performance was not predicted to rely on secondary memory, performance between the mismatch and no cue conditions did not differ. These results suggest that working memory relies on the same secondary memory component as episodic long-term memory tasks only under conditions predicted by a dual-component model of working memory (Unsworth & Engle, 2007).

Introduction

Working memory is the limited-capacity system responsible for the active maintenance and online manipulation of information over short periods of time (e.g., Baddeley, 1986; 2000). One reason why working memory has received considerable attention is because measures of working memory reliably predict individuals' performance on tasks measuring higher-order abilities. Indeed, performance on working memory tasks has been successful in predicting reading comprehension (e.g., Daneman & Carpenter, 1980), fluid intelligence (e.g., Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002), reasoning (e.g., Kyllonen & Christal, 1990), and complex learning (e.g., Shute, 1991; Tamez, Myerson, & Hale, 2012). The ubiquitous role of working memory in complex cognition have led some to consider it to be "the hub of cognition" (Haberlandt, 1997, p. 212), and "perhaps the most significant achievement of human mental evolution" (Goldman-Rakic, 1992, p. 111).

A Brief Overview of Working Memory

One important, early conceptualization of memory was proposed by Atkinson and Shiffrin (1968; see Figure 1), whose multi-store model posited separate systems for short-term memory (their version of working memory) and long-term memory. This non-unitary view of memory posited that short-term memory was critical for acquisition of long-term knowledge, and that more specifically, rehearsal of information in short-term memory was assumed to be necessary for transfer into long-term memory. However, the ability of individuals such as Patient KF (Shallice & Warrington, 1970; Warrington & Shallice, 1969) to acquire new long-term information despite an impaired short-term memory system was problematic for Atkinson and Shiffrin's theory, and called for alternative ways to conceptualize short-term memory.

Figure 1. Depiction of Atkinson and Shiffrin's (1968) multi-store model of memory.

As there were few individuals with Patient KF's specific pattern of cognitive deficits (impaired short-term memory with intact long-term memory), Baddeley and Hitch (1974) attempted to simulate pure short-term memory deficits in normal subjects using a dual-task technique. If performance on a digit span task (a task where participants are given a list of numbers that they must immediately repeat back in correct order) relies on short-term memory, it should be possible to occupy this store by varying the number of digits that need to be held in short-term memory, with larger numbers of digits occupying more of the store. According to Atkinson and Shiffrin (1968), increasing short-term memory load should systematically affect performance on other cognitive tasks that require transfer of information into long-term memory, eventually leading to a breakdown once the digit span load reaches the capacity of short-term memory. In other words, a near-capacity digit load should have damaging effects on a concurrent task that is dependent on long-term memory (e.g., reasoning, problem solving, comprehension, and learning). Baddeley and Hitch (1974) found, however, that the predicted detrimental effect of a near-capacity digit load on reasoning and problem solving was minimal. This in turn suggested that, by itself, the ability to store information cannot be responsible for higher-order abilities such as reasoning, comprehension, and learning. Instead, the notion of short-term memory needed to be expanded to include a more dynamic process beyond just storage.

This led to Baddeley and Hitch (1974) to propose what is perhaps the most well-known model of working memory: A multi-component model that is comprised not only of storage components (the phonological loop and visual-spatial sketchpad), but also includes a processing component (the central executive). This executive component is presumed to be responsible for a wide range of functions, including directing of attention to relevant information, inhibiting of irrelevant information and/or actions, and coordinating cognitive processes when more than one task must be done at the same time (see Figure 2).

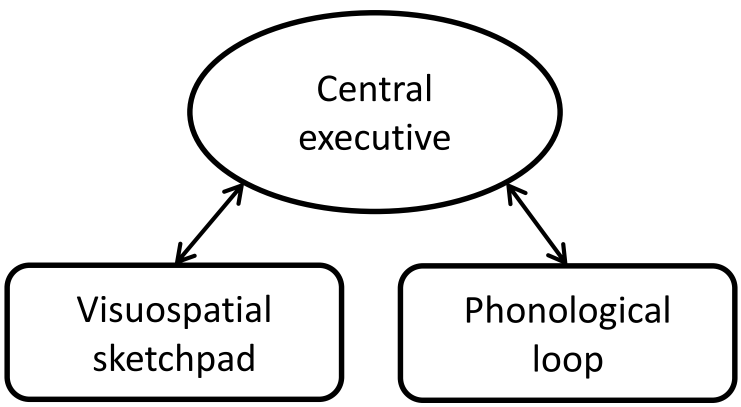


Figure 2. Depiction of Baddeley and Hitch's (1974) multi-component model of working memory.

Baddeley and Hitch's model of working memory was successful in explaining prior findings such as the word length effect (Baddeley, Thomson, & Buchanan, 1975) and the irrelevant speech effect (Colle & Welsh, 1976), as well as facilitating additional research. Since then, multiple models of working memory have emerged, each varying with regards to the structure of working memory, the contribution of long-term memory to working memory, the nature of working memory's capacity, and the role of attention in working memory efficiency (for an overview, see Miyake & Shah, 1999). However, central to just about all of these models of working memory are two ideas: 1) The "capacity" of working memory is limited, and 2) working memory is important for various aspects of higher order cognition.

Working Memory Tasks

As theories of working memory have expanded to include functions beyond just short-term memory, so have the tasks that are used to measure working memory. Early attempts to quantify working memory capacity were done using simple span tasks (e.g., digit span). However, such tasks did not seem to engage all of the components of Baddeley and Hitch's (1974) working memory model. In response, Daneman and Carpenter (1980) developed a task designed to tap the ability to both store and process information in short-term memory (i.e., complex span task). In their reading span task, participants are presented a series of unrelated sentences one at a time, which they are asked to read. At the end of the series, participants are asked to recall the last word from each sentence. Not surprisingly, as the number of sentences in the series increases, the more demanding the task becomes, as participants are forced to remember a greater amount of information while reading the sentences. Daneman and Carpenter (1980) found that performance on this task was a better predictor of reading comprehension than simple span tasks (see also Ackerman, Beier, & Boyle, 2005; Daneman & Merikle, 1996; Engle, Tuholski, Laughlin, & Conway, 1999). Since then, multiple complex span tasks have since been created that, although using different stimuli, preserve the general cognitive demands (i.e., storage and processing) that the reading span task imposes, and these complex span tasks have been used extensively to measure working memory ability (e.g., the operation span task; Turner & Engle, 1989).

Performance on complex span tasks has been found to be a significant predictor not only of performance on measures of theoretical constructs such as fluid intelligence (e.g., Engle et al., 1999; Süß et al., 2002), but of relevant classroom skills such as note-taking (e.g., Bui & Myerson, 2014), and even real-life outcomes such as performance on national curriculum assessments in young children (e.g., Gathercole, Pickering, Knight, & Stegmann, 2004). Furthermore, McCabe, Roediger, McDaniel, Balota, and Hambrick (2010) demonstrated that performance on complex span tasks is strongly correlated with measures of executive functioning, suggesting again that working memory is related to the control of complex cognition. The broad involvement of working memory has resulted in a great deal of attention with regards to the exact mechanisms that underlie it, as well as to why it successfully predicts a number of higher-order abilities.

Terminology

Before proceeding, it is useful to clarify terms that, although often used interchangeably, can be a source of confusion when discussing working memory. It is not uncommon to see the term primary memory used interchangeably with short-term memory, and secondary memory with long-term memory. However, Craik and Lockhart (1972) suggested that we should instead think about primary and secondary memory as systems, and short- and long-term memory as referring to the tasks/procedures that measure the contribution of those systems. Thus, interchangeably using these terms may not be semantically appropriate. Moreover, it may not be theoretically appropriate, as it implies that short-term memory tasks always measure primary memory, and that long-term memory tasks always measure secondary memory. As Watkins

(1977) has pointed out, the former may not be entirely accurate. With this in mind, these terms from here on out will be used following the systems/procedures distinction suggested by Craik and Lockhart.

The Role of Secondary Memory in Working Memory

Although early models of memory attempted to separate working memory and secondary memory (e.g., Atkinson & Shiffrin, 1968), studies have shown that this distinction fails to account for why maintaining familiar information is easier than maintaining novel information on a short-term memory task (e.g., Miller, Bruner, & Postman, 1954; Multhaup, Balota, & Cowan, 1996), or why an individual's memory span is typically smaller for unrelated words than for words that form a coherent sentence (e.g., Brener, 1940). Such examples suggest an influence of secondary memory in working memory, and have compelled psychologists to increasingly acknowledge the close relation between these two constructs (e.g., Jonides et al., 2008).

Several models have emerged specifying the relationship between working memory and secondary memory (e.g., Cowan, 1999; Oberauer, 2002; Unsworth & Engle, 2007a, 2007c). Indeed, Baddeley (2000) modified his original model by adding a new component, the episodic buffer, to accommodate the way in which working memory and secondary memory interact. Cowan (1999) proposed an embedded-process model suggesting that working memory is actually an activated subset (i.e., the "focus of attention") of secondary memory with a capacity of four chunks of information, and that all other items are retrieved from secondary memory. Similarly, Oberauer (2002) put forth a model in which information in memory exists in different states of accessibility, with only one chunk of information being in a state of direct access, while other information may remain in a less active state of readiness in secondary memory.

This present investigation is motivated by the dual-component model of working memory put forth by Unsworth and Engle (2007a, 2007c). According to their model, working memory partly reflects the ability to retrieve information from secondary memory when to-be-remembered information is displaced from primary memory. They assume that there is a limit on the amount of information that can be actively maintained in primary memory at any given time, which, following Cowan (1999), they posit is about four chunks (for similar estimates, see also Watkins, 1974). When this limit is exceeded, some of the to-be-remembered information is displaced into secondary memory, from which it can be retrieved later. Unsworth and Engle have suggested that to-be-remembered information also may be displaced into secondary memory if the attentional resources needed to maintain the information in primary memory are diverted (i.e., a secondary task).

These ideas suggest that commonly used short-term memory tasks engage primary memory and secondary memory to different extents. For example, simple span tasks such as the digit span task presumably measure the ability to maintain a list of items and report them directly from primary memory, unless the list length exceeds the capacity of primary memory, at which point both primary and secondary memory are involved (e.g., Unsworth & Engle, 2007a, 2007c; Watkins, 1977). In contrast, complex span tasks such as the reading span task or the operation span task require participants to perform a secondary processing task that is interleaved between the presentations of the to-be-remembered items. According to Unsworth and Engle's model, such secondary tasks require that participants temporarily switch attention away from maintaining the to-be remembered items in primary memory, meaning that although some items may be reported from primary memory (such as the last item in the series), others must be retrieved from secondary memory. In short, this model suggests that both primary and secondary memory are involved in performing simple and complex span tasks, although the former relies more on primary memory, while the latter relies more on secondary memory.

Unsworth and Engle's (2007a, 2007c) working memory model posits that under certain circumstances, simple span tasks can be just as good a measure of working memory as complex span tasks. More specifically, if simple span tasks with longer list lengths rely on secondary memory in a manner similar to complex span tasks of any length, then both types of task are assumed to be engaging processes critical for working memory (cue dependent search from secondary memory). This suggestion is important considering previous findings suggesting that simple span tasks and complex span tasks tap into separate constructs that are differentially related to higher-order cognitive abilities, with the former measuring short-term memory and the latter measuring working memory (e.g., Ackerman, Beier, & Boyle, 2005; Daneman & Merikle, 1996; Engle et al. 1999).

However, it may not necessarily be the case that simple and complex span tasks always tap into separate constructs. For example, results from Unsworth and Engle (2006) showed that the correlation between performance on a complex span task and fluid intelligence task does not systematically change as a function of series length, whereas the correlation between simple span task and fluid intelligence increases as series length increases, eventually reaching (after a series length of four) similar magnitudes of correlation as complex span task and fluid intelligence. These findings are consistent with a dual component model of working memory (Unsworth & Engle, 2007a, 2007c), in that if performance on complex span tasks generally relies on secondary memory, its correlation with fluid intelligence should not vary as a function of series length (see also Salthouse & Pink, 2008). Conversely, if performance on a simple span task only relies on secondary memory if the series length is greater than four, then no correlation with fluid intelligence should be observed on subspan series lengths (2-4 items), whereas a correlation with fluid intelligence should be observed on supraspan series lengths (5-7).

Secondary Memory in Working Memory and Long-Term Memory Tasks

Broadly speaking, the primary goal of the present investigation is to answer the question: Is the secondary memory component involved in working memory the same as the secondary memory component in long-term episodic memory? Indeed, the answer to this question may allow examination of whether the governing dynamics of long-term episodic memory also apply to working memory. To the extent that both long-term episodic memory and working memory rely on the same secondary memory component, it stands to reason that experimental manipulations should produce patterns of effects on working memory tasks similar to those observed on long-term episodic memory tasks. This was the approach taken by Rose, Myerson, Roediger, and Hale (2010), who examined the effects of the level of processing of to-be-remembered items (Craik & Tulving, 1975) on a working memory task. In this study, Rose et al. (2010) presented participants with a target word (*"bride"*, presented in red), which was followed by two "processing" words ("*dried"*, presented in blue, and "*groom"*, presented in red). Additionally, the study was constructed such that the target word could be matched with one of the two processing words based on the color of the word (shallow, visual processing), rhyme (intermediate, phonological processing), or meaning (deep, semantic processing).

After being presented with several target words, participants were given a recall test. The results of this study demonstrated that attending to different types of features (visual, phonological, semantic) of words at the time of encoding did not produce effects on working memory tasks like those typically observed on long-term memory tasks (i.e., retention being better for semantically processed items than for phonological and visual processing; Craik &

Tulving, 1975). However, subsequent studies have shown such effects of level of processing on working memory tasks (Loaiza, McCabe, Youngblood, Rose, & Myerson, 2011; Rose & Craik, 2012; Rose, Buchbaum, & Craik, 2014), prompting Rose and Craik (2012) to suggest that working memory tasks and long-term episodic memory tasks may both tap into similar secondary memory systems.

This approach of modifying working memory tasks to incorporate manipulations known to produce robust effects on long-term episodic memory tasks provides a unique opportunity to study the relation between the secondary memory components thought to underlie performance on each task. Indeed, the evidence reviewed in this section suggests that manipulations of encoding condition yield similar effects on working memory tasks and long-term episodic memory tests. However, it is unclear whether manipulation of retrieval conditions would yield similar results. This begs the question of whether manipulations of retrieval conditions known to affect long-term episodic memory would have similar effects on working memory. This is the central question addressed in this investigation, but before the current set of studies are described in more detail, some important aspects of retrieval on long-term episodic memory tasks that will guide the predictions made by this investigation are provided in the following section.

Encoding Specificity Principle

One early assumption about memory was that retrieval of an event, as measured by responses on a memory test, reflected only the contents of what was stored in memory, and any failure in retrieval was attributed to failures in encoding or storage. Although Tulving (1974) acknowledged that the encoding process was important in understanding memory, he also emphasized that memory for an event depended on the cues present at the time of retrieval. Because multiple pieces of information may be linked to a single cue and one piece of information may be linked to multiple cues (e.g., Watkins & Watkins, 1975), conditions under which retrieval is successful depend on the context in which it was initially encoded. Simply put, retrieval of information (and thus forgetting, as operationally defined) is cue dependent (e.g., Tulving, 1974), and therefore failure to retrieve target information may not always reflect unavailability of that information, but rather may sometimes reflect the inaccessibility of that information due to the presence of inappropriate cues at the time of the retrieval attempt (see Tulving & Pearlstone, 1966).

The cue-dependent nature of memory retrieval is highlighted in a study by Thomson and Tulving (1970; Experiment 2), who presented participants with a list of weakly related word-pairs ("train : BLACK") to study. Participants were told that they would need to remember the capitalized target words ("BLACK") for a later test, but that they should also note its relation to the first word ("train"). At the time of the test, participants either got the cues that were the same as the ones shown during study ("train"; match condition), cues that were strongly associated with the target word but were not shown during study ("white"; mismatch condition), or no cues at all (no cue condition).

What Thompson and Tulving found (see Figure 3) was that giving participants the mismatch cues (M = .23) led to a lower proportion of items being recalled than giving them the match cues (M = .83), despite that fact that the mismatch cues were more strongly associated with the target word than the match cues. Furthermore, and perhaps more interesting, mismatch cues led to worse performance than in the no cue condition (M = .30). Although this difference was not statistically significant in Thomson and Tulving's study, later studies found significant differences, providing even stronger evidence for the encoding specificity principle (e.g., Roediger & Adelson, 1980; Roediger & Payne, 1983; cf. Newman & Frith, 1977). Specifically, these studies provide evidence that no matter how strongly a cue is associated with the target word, it will not be effective in facilitating recall unless the target was initially encoded with respect to that specific cue.

Further support for the encoding specificity principle is provided by recognition failure of recallable words. In one example given by Tulving and Thomson (1973), participants were presented a list of weakly related word-pairs ("train : BLACK") and were instructed to remember the second word ("BLACK") for a later test. Afterwards, participants were given a list of words

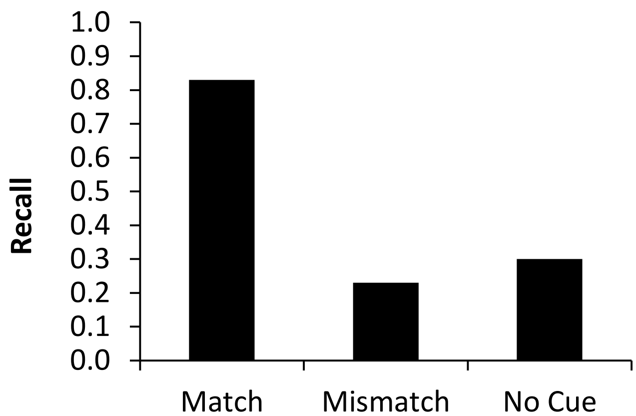


Figure 3. Partial results from Thomson and Tulving's Experiment 2 (1970) depicting recall on the final trial as a function of cue condition.

("white"), to which they were asked to generate six associates. In this example, participants are likely to generate the word "black", in addition to other associates. Of these generated associates, participants were then asked to circle the ones that had appeared on the initial list of target words to be learned. Interestingly, participants are quite poor at correctly recognizing the target words ("black") that they themselves had generated, doing so only 24% of the time. However, when participants were later given the original cue ("train"), and asked to recall the target word ("BLACK"), they were successful at doing so 63% of the time. Thus, "black" is initially encoded in the context of "train" and is not recognized as having been presented earlier when generated in the context of “white”. The results of this study demonstrate that despite being able to recall the target words when given the original cues, participants are unable to recognize those same words in the context of a new cue, even when they generated the word themselves from a new cue (see also Nilsson, Law, & Tulving, 1988).

Finally, more subtle effects of encoding specificity can also be observed, in which if a certain characteristic of an object is emphasized at encoding, then other salient aspects of that object will not be as effective at the time of retrieval. For example, Barclay, Bransford, Franks, McCarrell, & Nitsch (1974) had participants study sentences such as “The man lifted the piano” or “The man tuned the piano”. Results indicated that recall using a cue such as “something heavy” was better for the first sentence, whereas a cue such as “a musical instrument” was better at promoting recall of the second sentence, demonstrating that successful retrieval can depend on more understated aspects of the similarity of encoding and retrieval operations.

The encoding specificity principle rejects the idea that the general strength of the association between the cue and target is the primary determinant of whether the target word is recalled, and instead posits that under specific circumstances, even weakly associated cues can be more effective in facilitating recall compared to a strongly associated cue. Furthermore, cues that were present at the time of test, but which were not present at the time of study, can lead to worse recall than having no cues at all. The common assumption at the time, as exemplified by Anderson and Bower (1972) was that any cue with a prior relationship to the target word should facilitate recall, and that strongly associated cues should help recall more than weakly associated cues, and any cues should help recall more than having no cues at all. As Thomson and Tulving (1970) showed, however, this is simply not the case, and the scenarios described in the previous paragraphs highlight instances where generate-recognize theories fail. In short, the encoding specificity principle drastically changed the way that researchers subsequently approached memory research by placing more emphasis on both the powerful and the more subtle effects of cues on memory.

On a broad level, the encoding specificity principle is the idea that retrieval of information from secondary memory is generally more successful when the cues present at retrieval are also present at the time of encoding, and that cues affect the likelihood of retrieval in both positive and negative ways. However, although these principles have been shown to hold for long-term episodic memory tests, it is not yet known whether this type of cue dependency also holds for short-term memory tests thought to measure working memory. If it does, this would provide converging evidence that the same secondary memory system is involved in performance on working memory and long-term episodic memory tasks. As already noted, this question represents the main focus of the present investigation. But before describing it in further detail, it is important to consider primary memory in the context of the assumptions made by the encoding specificity principle.

The Nature of Primary Memory

One important observation that arise from the earlier discussed studies examining encoding specificity is that the associations made between the cue and the target word are semantic in nature. Indeed, manipulating the cue-target pairings based on pre-experimentally established associations was key in refuting the assumptions of the generate-recognize models, though as Tulving (1983) points out, the binding between cue and target do not necessarily have to be semantic in nature. Along with the fact that these studies assessed performance using long-term memory tests, which presumably relies exclusively on secondary memory, it is reasonable to assume that items in secondary memory are semantically coded. But how are items stored in primary memory coded? This is important to consider because this investigation examines whether encoding specificity effects are observed when performing short-term memory tasks, which have been argued to rely on secondary memory only under certain conditions (Unsworth & Engle, 2007; 2007c). But to the extent that performance on short-term memory tasks rely on primary memory, does the manner in which information in primary memory is coded influence whether encoding specificity is observed?

The prevalent view early on was that information stored in primary memory was phonologically coded, whereas information stored in secondary memory were semantically coded. This view was supported by studies that manipulated the phonological and/or semantic similarity between to-be-remembered words, and examining its effect on recall. For example, Baddeley and Dale (1966) presented participants with two lists that each contained eight word-pairs, which was then followed by a test on the first list. In a control condition, the two lists were not related, whereas in the experimental condition, the two lists each contained word-pairs that were semantically similar to each other. In this paradigm, participants in the control condition performed better on the first list compared to participants in the experimental condition. In the short-term memory version of this experiment, Baddeley and Dale presented participants with a short list of word-pairs, which they were tested on a few seconds later. Again, in the control condition, the word-pairs in the list were not semantically similar, whereas in the experimental condition, they were. Results indicated that semantic similarity did not have an effect on recall, suggesting that items in primary memory were not semantically coded.

Similarly, Baddeley (1966) presented participants a list of ten items that were either high in phonological similarity or high in semantic similarity. At immediate recall (when performance is presumed to rely on primary memory), phonological similarity was harmful for performance, whereas semantic similarity was not. Conversely, when a 30-second distractor task was interleaved between presentation and recall (where performance presumably relied on secondary memory), the opposite was observed: Performance was hurt by semantic similarity, but not by phonological similarity.

Similar divisions suggesting that primary and secondary memory are organized according to different codes can also be found with studies examining the serial position curve. In particular, one robust finding in free recall is that memory is usually best for items that were presented towards the end of the list compared to items in the middle of the list (recency effect). One explanation has been that the recency effect reflects items that are stored in primary memory, whereas the middle of the serial position curve reflects items contained in secondary memory (e.g., Waugh & Norman, 1965; cf. Bjork & Whitten, 1974). To the extent that this is the case, we may expect to see phonological and semantic similarity have differential effects on particular portions of the serial position curve. To demonstrate this, Kintsch and Buschke (1969) presented participants with a list of 16 words, after which a single probe word from that list was shown again, and participants were asked to recall the word that came after the probe. In one condition, all the words were unrelated, whereas the other condition contained eight pairs of synonyms scattered throughout the list. The list containing synonyms led to worse recall compared to the list of unrelated words, though this was only observed for items in the prerecency portion of the serial position curve. Conversely, when the eight pairs of synonyms were replaced with homonyms, the opposite pattern was observed: Recall was hurt for items in the recency portion of the serial position curve.

This simplified division, however, has not always received empirical support. For example, Shulman (1970) presented participants with a list of words, which was followed by a yes-no question regarding a probe word from the studied list. Specifically, these questions asked whether the probe word was either identical to any of the words from the study list, a homonym of one of the words from the study list, or a synonym of one of the words from the study list. For the present purposes, we will focus on the synonym questions, which Shulman found that performance was best for items in recency portion of the serial position curve. Shulman took these results to suggest that primary memory do in fact contain semantic codes. However, Baddeley (1972) has suggested an alternative explanation for Shulman's results, pointing out that participants may be generating semantic associates to the probe word, and then comparing the phonological representation of those associates to the list items. In this scenario, semantic codes are never actually carried in primary memory, but are instead produced by the participants when they answer the yes-no questions. Though this is indeed possible, it should be noted that this not-so-parsimonious explanation likely requires a rapid set of processes within a very short period of time.

Finally, a study by Watkins, Watkins, and Crowder (1974) presented participants with a list of words that were either phonologically similar to one another, or phonologically dissimilar. Afterwards, participants were asked to free recall the words. Phonological similarity did not have an effect on items in the recency portion of the serial position curve, but did help memory for items in the prerecency portion of the curve. The results of this study suggests that phonological codes are stored secondary memory, which is actually consistent with more recent findings demonstrating that when using a Deese/Roediger-McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) paradigm, phonological false memories can occur (e.g., Sommers & Lewis, 1999). As a result, it may be the case that phonological codes are in fact stored in secondary memory, and it might not be necessarily be the case that information is translated from phonological to semantic codes when information is transferred from primary to secondary memory, but that the information in secondary memory is tagged with semantic information as well.

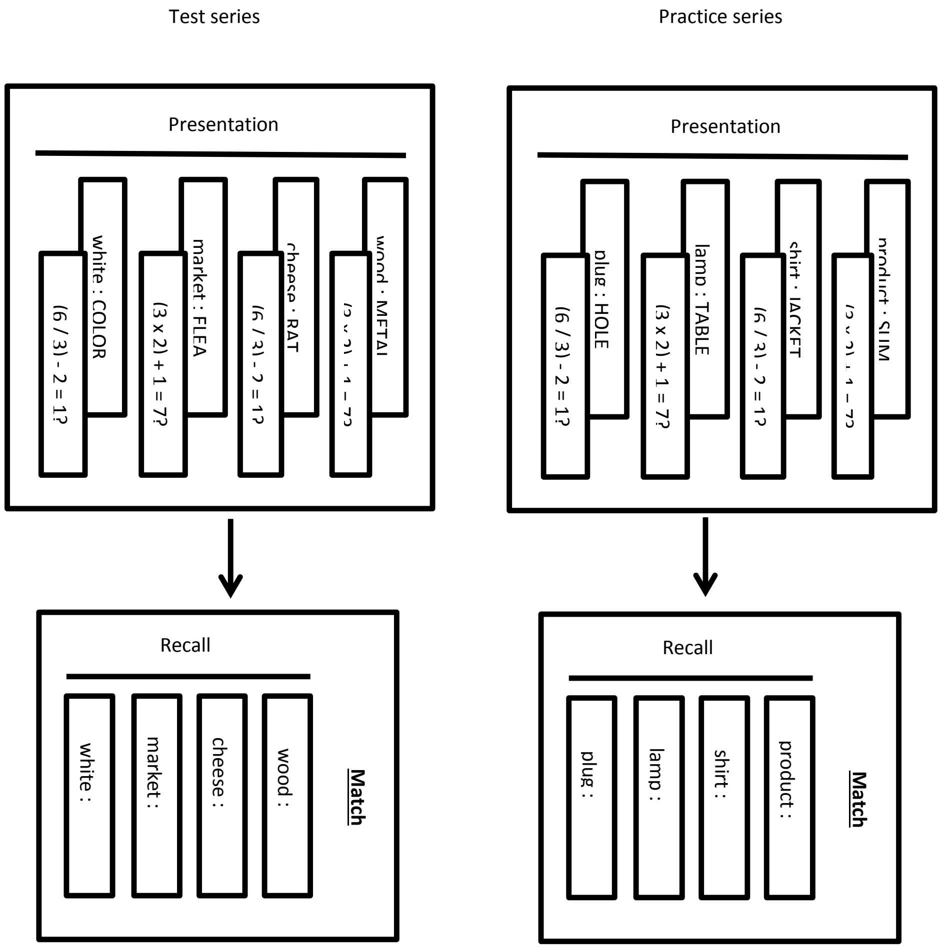
For the purposes of the current investigation, it should be noted that the studies discussed so far do not provide any evidence that information in primary memory is semantically coded. If recall on a short-term memory test does rely on secondary memory, it would be reasonable to believe that performance would be worse with mismatch retrieval cues compared to recall with no cues (e.g., Thomson & Tulving, 1970). However, to the extent that primary memory does not contain semantic codes, whether this pattern would be observed if performance only relied on primary memory is less clear.

Current Study

The premise of the current investigation is that knowledge about the role of cues in retrieval in long-term episodic memory can inform the understanding of retrieval processes that take place in working memory. Unsworth and Engle's (2007a, 2007c) model of working memory makes some interesting predictions when considering the hypothesized relation between working memory and secondary memory. According to their model, items are displaced from primary memory into secondary memory under two conditions: When the capacity of primary memory is exceeded (such as in a simple span task with a longer series lengths), and when attentional resources needed to maintain the to-be-remembered information are diverted to a secondary task (such as in a complex span task). Under these conditions, working memory is thought to rely on retrieval from secondary memory. Given the literature described earlier, to the extent that both working memory and long-term episodic memory share overlapping processes, it would seem to follow that the type of cues present at retrieval should affect performance on a working memory task in the same way that it affects long-term episodic memory tests.

Therefore, the first goal of this investigation was to apply the encoding specificity principle to short-term memory tasks that measure working memory, and to answer the question: Is performance on a short-term memory task cue-dependent, such that mismatching retrieval cues at the time of test produces worse recall than either receiving matching cues, or no cues at all? If the encoding specificity principle, which describes performance on long-term episodic memory tests, also describes performance on working memory tasks, this would support the hypothesis that the secondary memory component is also the same as the one underlying long-term episodic memory.

The current study consisted of three experiments, each using an approach analogous to that used previously by Rose and colleagues (Rose et al., 2010; Loaiza et al., 2011; Rose & Craik, 2012) that involves modifying a short-term memory task to include manipulations known to have robust effects on long-term episodic memory tests. As shown in Figure 4, an operation span task was altered to include weakly related word-pairs ("era : TIME") as the to-be-remembered items instead of single words, which are usually used as the to-be-remembered items (e.g., Turner & Engle, 1989). At the end of each series, participants were asked to recall the target words ("TIME"). On recall after the first four series, participants were provided with the same cue as from the study phase ("era"; match condition). In order to test the encoding specificity principle, on the final (fifth) series at recall participants were provided with either a cue that was the same as from the study phase (match condition), a strongly related cue different from what was presented during the study phase ("life"; mismatch condition), or were asked to free recall the target words in any order (no cue condition).



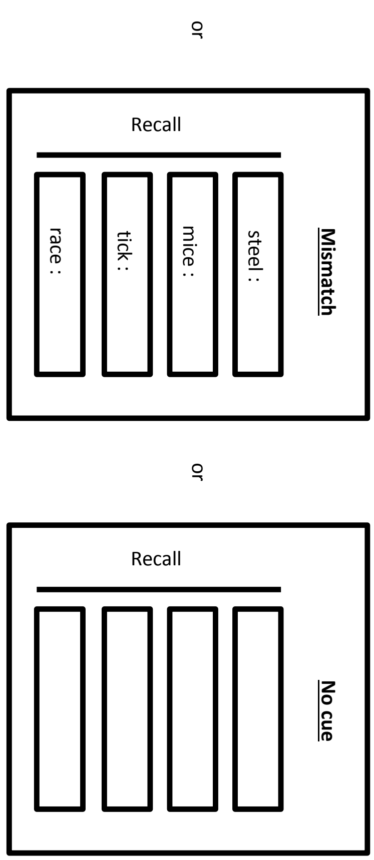


Figure 4. The general procedure used in the reported experiments. Depicted is an example of modified operation span task for a 4- item series. Participants completed four practice series where, during the recall phase, they were given the same cue as the one given during the presentation phase. Afterwards during the recall phase of the test series, participants were either given the same cue as the one during presentation (match), a different but related cue (mismatch), or told to free recall the target words (no cue).

This procedure was used for two reasons. As in Thomson and Tulving (1970): 1) The initial practice series were intended to allow participants to develop confidence that the cues at the time of study would be presented during test, and thus encourage them to form semantic associations between the cue and target; 2) after the final (fifth) series, participants in the mismatch condition would be less likely to make any semantic associations between the cue and target, making interpretation of recall during any subsequent series difficult. To remain consistent with scoring procedures used by Thomson and Tulving (1970), accuracy after each series was measured as the proportion of total targets recalled, regardless of order. If recall on a working memory task is in fact cue dependent, then one would expect recall in match conditions to be better than recall in both the mismatch and no cue conditions. Perhaps more importantly, one would expect to see recall for the mismatch condition to be worse than the no cue condition.

The second goal of this investigation was to examine the boundary conditions under which encoding specificity effects are observed in short-term memory tests. As mentioned earlier, Unsworth and Engle (2007a, 2007c) have suggested that series length and secondary task demands are two factors that influence when performance on a short-term memory test relies on secondary memory. It follows, then, that experimentally manipulating these two factors should provide an excellent opportunity to create conditions under which performance on a short-term memory task is predicted to rely on secondary memory. In the case of the current investigation, systematically examining when performance on a short-term memory task is cue dependent allows one to test the predictions based on the dual component model.

To address this second goal, the present experiments varied series length and secondary task demands. More specifically, the two series lengths used were either subspan (4 items) or supraspan (8 items), and secondary task demands were either present (making it a complex span task) or not present (making it a simple span task). As can be seen in Figure 5, this allowed examination of four possible combinations of when performance on a short-term memory task may rely on secondary memory.

		Length	
		Subspan	**Supraspan**
Task	**Complex**	Secondary Memory	Secondary Memory
	Simple	Primary Memory	Secondary Memory

Figure 5. A matrix depicting the four possible combinations of when performance on a working memory task is hypothesized to rely on secondary memory. Experiment 1 examined a complex span task with supraspan list lengths; Experiment 2 examined a simple span task with supraspan list lengths; Experiment 3 examined both complex and simple span tasks with subspan list lengths.

In scenarios where performance does rely of secondary memory, one would expect to see patterns supporting the encoding specificity principle (match > no cue > mismatch). Moreover, these patterns should be observed when either the capacity of primary memory is exceeded, or when there are secondary task demands present. One would expect this pattern for the cells in Figure 5 corresponding with complex/subspan-length, complex/supraspan-length, or simple/supraspan-length, but not for the cell corresponding with simple/subspan-length.

In summary, all three experiments examined whether performance on a short-term memory task was cue dependent, in accordance with the encoding specificity principle. As mentioned above, each experiment examined this possibility as a function of series length and secondary task demands, which allows for the testing of whether encoding specificity is observed under conditions predicted by Unsworth and Engle (2007a, 2007c). More specifically, performance was examined in Experiment 1 under complex/supraspan-length conditions, in Experiment 2 under simple/supraspan-length conditions, and in Experiment 3 under both complex/subspan-length and simple/subspan-length conditions.

Experiment 1 (Complex/Length 8)

In this experiment, participants performed a modified operation span task with a series length of eight items. According to Unsworth and Engle (2007a, 2007c), participants should have to rely on secondary memory in this situation not only because the capacity of primary memory is exceeded, but also because the demands of the arithmetic problems force the contents of primary memory to be displaced into secondary memory. Thus, if the encoding specificity principle applied to recall on this modified span tasks, it would be predicted that performance in the match condition would be better than both the mismatch and no cue conditions, and that the mismatch condition would be worse compared to the no cue condition.

Method

Participant and Design

A total of 75 participants (51 females, 24 males; $M_{\text{age}} = 36.4$ years, $SD = 12.9$) were recruited from the pool of workers maintained by Amazon Mechanical Turk (MTurk)[1] and

[1] MTurk has been proven to be a reliable method for collecting cognitive data, as evidenced by internet-based studies replicating laboratory findings of effects such as the Deese/Roediger-McDermott false memory effect (Bui, Friedman, McDonough, & Castel, 2013), the spacing effect (Bui, Maddox, & Balota, 2013), the levels of processing

received monetary compensation ($0.40) for their participation. All participants reported proficiency in English and resided in the United States. Final test cue (match vs. mismatch vs. no cue) was manipulated between participants, with 25 participants assigned to each condition.

Materials

Five series, each consisting of eight word-pairs, were used for all participants. Forty target words were chosen (see Appendix A), each with a strongly associated cue (FSG range: 0.25-0.32) as well as a weakly associated cue (FSG range: 0.01-0.02), from a database of word associates (Nelson, McEvoy, & Schrieber, 1998). Across all target words, backward association strength was held constant, and each cue word was associated only with its intended target word and not with any of the other 39 target words as determined by the Nelson et al. database.

In addition to the word-pairs, 40 arithmetic problems ("[4 x 2] - 1 = 9?") were chosen for the operation span task. These arithmetic problems were selected from an unpublished database containing approximately 300 arithmetic problems that have been normed across 294 participants of various ages (18-70 years). The arithmetic problems were selected such that they were similar with regards to how long they took on average to verify, which reduced the amount of variability in the duration of the secondary task (arithmetic problems).

Procedure

Participants were told that they would see a series of word-pairs ("era : TIME") alternating with arithmetic problems, and that they would need to remember the word-pairs for a later test. The cues paired with target word during presentation were always weakly associated cues. In addition, participants were told that during the test, they would be presented with the first word ("era"; cue), and that they would have to recall the target word ("TIME") associated

---

effect (Bui, Maddox, Zou, & Hale, 2014), and age-related declines in processing speed and working memory (Bui, Myerson, & Hale, under review).

with that cue. They were also told that this cue would either be the same as the one presented earlier during the series (“era”; match cue), or a different cue not previously presented (“life”; mismatch cue). Additionally, participants were told that although the mismatch cues were not presented earlier, they were strongly associated with the target word, and that participants should use that cue to aid recall on the test. The exact set of instructions given to participants is provided in Appendix B.

Each series began with a word-pair being presented on the screen for 3 seconds, after which an arithmetic problem was displayed. Participants verified whether the arithmetic problem was correct or incorrect by clicking a button on the screen. At the end of each series, participants were given a single cue at a time, and asked to recall the target word that was paired with it (or in the mismatch case, the target word related to the cue). Participants typed in their response, and hit the Enter key to move on to the next cue. Participants were given up to ten seconds to recall each target word, and the order of the cues presented on the test was the same as that during the earlier presentation.

During recall on the first four series, the cue presented was always the same as that used during the earlier presentation (match). On recall during the fifth series, participants were either given the same cues (match), a different cue that is strongly associated with the target word (mismatch), or were asked to free recall the target words in any order they wanted (no cues).

Results and Discussion

The mean response time on the arithmetic problems was 3874.58 ms, and the mean proportion of correct responses to the problems was .84. Moreover, removing participants who responded correctly to the arithmetic problems less than 50% of the time did not change any of the pattern of results reported below. Figure 6 presents accuracy of recall on the test series as a function of cue condition (i.e., match, mismatch, and no cue). A lenient scoring method was taken from Thomson and Tulving (1970), such that for all three conditions, responses were scored independent of serial order. For example, participants received credit even if the first word recalled was actually the third word that was presented earlier during the series. A one-way analysis of variance (ANOVA) revealed no effect of cue condition on recall on the practice series, $F < 1.00$. However, there was an effect of cue condition on recall on the test series, $F(2, 72) = 34.56$, $p < .001$, $\eta_p^2 = 0.49$.

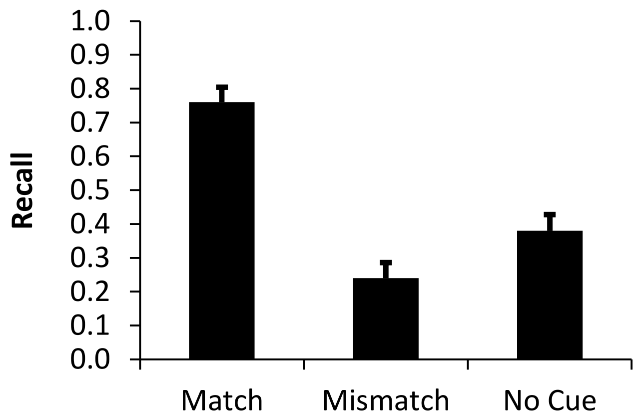


Figure 6. Performance on test series as a function of cue condition in Experiment 1. Error bars represent standard error of the mean.

Follow-up comparisons indicated that recall in the match condition was significantly better than in either the mismatch condition ($p < .001$, $d = 2.31$) or the no cue condition ($p < .001$, $d = 1.65$). Critically, recall on the test series was better for the no cue condition than recall in the mismatch condition, $p = .033$, $d = 0.60$. The results of Experiment 1 suggest that—similar to performance on long-term episodic memory tasks—accuracy of recall on a complex span task is cue dependent, and provides converging evidence that the secondary memory component underlying working memory is the same as that underlying long-term episodic memory.

However, it remains unclear as to whether the to-be-remembered items in this experiment were displaced into secondary memory because the capacity of primary memory (four items) was exceeded, or because the attention needed to maintain those items in primary memory was diverted to a secondary task (arithmetic problems), or both. Indeed, an operation span task with a series length of eight not only exceeds the capacity of primary memory, but presumably also requires diverting attention away from the to-be-remembered items when performing the arithmetic task. Thus, the implicit assumption by Unsworth and Engle (2007a, 2007c) that both methods of displacement produce similar encoding specificity effects is not necessarily supported by the results of this experiment. In order to validate such an assumption, recall on the modified span task must be observed when only secondary task demands are present, or when only the capacity of primary memory is exceeded.

Experiment 2 (Simple/Length 8)

The goal of Experiment 2 was to further investigate the role of secondary memory in working memory tasks, and to examine whether retrieval on the modified span task is cue dependent when to-be-remembered items are displaced into secondary memory due of capacity limitations. To do so, the word-pairs in Experiment 2 were presented without intervening arithmetic problems, which remove any secondary task demands. In such a simple span task with supraspan series length, any evidence of encoding specificity would presumably be due solely to the capacity of primary memory being exceeded. Accordingly, it is predicted that performance in the mismatch condition would be worse compared to the no cue condition.

Method

A total of 150 participants (103 females, 47 males; $M_{\text{age}} = 27.4$ years, $SD = 6.5$), were recruited from MTurk to take part in this study for monetary compensation ($0.40). All participants reported proficiency in English and resided in the United States. Final test cue (match vs. mismatch vs. no cue) was manipulated between participants, with 50 participants assigned to each condition. The materials and procedure for Experiment 2 were the same as Experiment 1, except that no arithmetic problems were given in between the presentation of the to-be-remembered word-pairs.

Results & Discussion

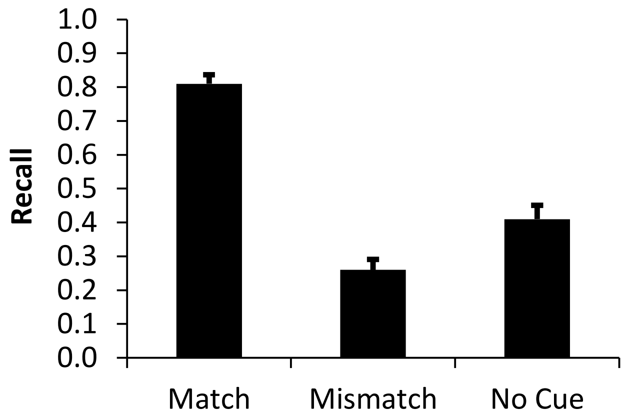


Figure 7. Performance on the test series as a function of cue condition in Experiment 2. Error bars represent standard error of the mean.

Recall on the test series as a function of cue condition is presented in Figure 7. A one-way ANOVA revealed no effect of cue condition on recall on the practice series, $F < 1.00$, though there was an effect of cue condition on recall on the test series, $F(2, 147) = 73.17$, $p < .001$, $\eta_p^2 = 0.50$. Follow-up comparisons indicated that subjects in the match condition correctly recalled more target words than both the mismatch condition ($p < .001$, $d = 2.68$) and no cue condition ($p < .001$, $d = 1.62$). More importantly, recall in the no cue condition was significantly better than recall in the mismatch condition ($p = .002$, $d = 0.58$).

The results of Experiment 2 provide further evidence that recall on a working memory task is cue dependent, and converging evidence that similar secondary memory systems underlie both working memory and long-term episodic memory. Further, these patterns were obtained using a simple span task, suggesting that secondary task demands do not have to be present in order for working memory to be involved. In Experiment 2, it is presumed that working memory demands were necessitated by the supraspan nature of the task. These patterns of results are similar to Experiment 1, although the underlying reasons may be quite different across the two experiments. In Experiment 1, the cue dependent nature of recall on the complex span task may have been due to either secondary task demands or the capacity of primary memory being exceeded. Contrary, in Experiment 2, secondary task demands were eliminated by using a simple span task, and in doing so, provides evidence that encoding specificity can be observed when only the capacity of primary memory being exceeded.

Experiment 3 (Complex vs. Simple/Length 4)

In Experiments 1 and 2, cue dependency was observed in recall on a working memory task where both secondary task demands are present and the capacity of primary memory is exceeded (Experiment 1), and when only capacity is exceeded (Experiment 2). Though it may be the case that the encoding specificity in Experiment 1 was due to secondary task demands, the supraspan nature of the task makes it difficult to say so with certainty. It is possible that secondary task demands do not actually displace to-be-remembered information into secondary memory, and instead the cue dependent recall observed in Experiment 1 are due to the capacity of primary memory being exceeded. To isolate the effects of secondary task demands on recall in a short-term memory task, it was necessary to use a complex span task whose list length did not exceed the capacity of primary memory. Furthermore, to establish that this pattern of encoding specificity (mismatch < no cue) only occurs with secondary task demands or when primary memory capacity is exceeded, it was important to demonstrate that when neither of those conditions are met, the pattern of results are different from what is observed if one or both of those conditions are met.

To do so, the series length on the modified span task was shortened to four, and the presence of intervening arithmetic problems was manipulated between subjects (complex vs. simple span). The subspan nature of the task allows for the direct examination of the role of encoding specificity in recall on a short-term memory task solely as a function of secondary task demands (complex vs simple). In the complex span condition, the presence of a secondary task should necessitate the use of secondary memory, and it is predicted that recall on mismatch series be worse than recall on the no cue series. In the simple span condition, the lack of secondary task demands (combined with a subspan series length) should mean that the to-be-remembered items are able to be maintained in primary memory. As secondary memory is not involved in this instance, it is predicted that recall in the mismatch condition be no different (or better) compared to recall in the no cue condition.

Method

Participants and Design

A total of 200 participants (109 females, 91 males; $M_{\text{age}} = 28.8$ years, $SD = 6.2$), were recruited from MTurk to take part in this study for monetary compensation ($0.40). All participants reported proficiency in English and resided in the United States. Final test cue (mismatch vs. no cue) and task demand (simple vs. complex) were both manipulated between participants, with 50 participants assigned to each condition. In Experiments 1 and 2, because recall on the practice series for all three cue conditions did not differ from one another, nor did they differ from recall on the test series for the match condition, the match condition was dropped in this experiment.

Materials and Procedure

The materials and procedure for Experiment 3 were the same as Experiment 1 and 2, except for two critical differences: 1) series length was four items, and 2) arithmetic problems were interleaved in between the presentation of the word-pairs for the complex span condition, but not for the simple span condition.

Results and Discussion

The mean response time on the arithmetic problems in the complex conditions was 4018.84 ms, and the mean proportion of correct responses to the problems was .82. As was the case in Experiment 1, removing participants who responded correctly to the arithmetic problems less than 50% of the time did not change any of the pattern of results reported below. Recall on the test series as a function of cue condition and task demands are presented in Figure 8. With regards to performance on practice series, a two-way ANOVA revealed that performance on simple span tasks ($M = 0.86$) was marginally better than complex span tasks ($M = 0.82$), $F(1,196) = 3.05$, $p = .082$, $\eta^2 = 0.02$. Additionally, there were no differences in performance between the mismatch condition ($M = 0.83$) and no cue condition ($M = 0.85$), nor was there an interaction between the two factors, F's < 1.00.

When examining performance on the test series, a two-way ANOVA revealed that recall on simple span tasks ($M = 0.53$) was better than recall on complex span tasks ($M = 0.37$), $F(1,196) = 16.69$, $p < .001$, $\eta^2 = 0.08$. There were no differences in recall between the mismatch condition ($M = 0.41$) and no cue condition ($M = 0.48$), $F(1,196) = 2.71$. However, there was a significant interaction between cue condition and task demand on recall accuracy, $F(1,196) = 4.16$, $p = .043$, $\eta^2 = 0.02$. Follow-up comparisons indicated that for complex span tasks, recall in the no cue condition was better than recall in the mismatch condition, $p = .014$, $d = 0.50$. More importantly, recall in the simple span tasks did not differ as a function of the two test cue conditions, $p = .754$, $d = 0.07$.

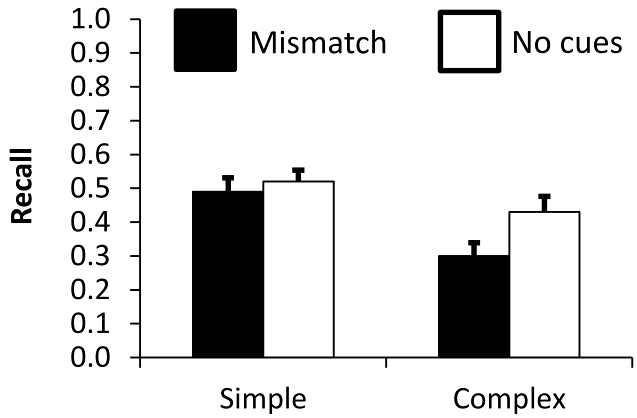


Figure 8. Performance on the test series as a function of cue condition and task type in Experiment 3. Error bars represent standard error of the mean.

The results of this experiment demonstrate two critical points: 1) encoding specificity (defined as mismatch < no cue) effects can be observed on a short-term memory task under conditions where secondary task demands elicit the need for secondary memory, but 2) not during recall of a subspan series length when there are no secondary task demands. Importantly, this experiment allows examination of the effect of secondary task demands on recall with lessened concerns that the capacity of primary memory was exceeded, a concern that limited interpretation of the results from Experiment 1. Theoretically, the results from Experiment 3 lend further support to Unsworth and Engle's (2007a, 2007c) suggestion that secondary task demands displace to-be-remembered information from primary memory into secondary memory. Perhaps more critically, these results provide evidence for Unsworth and Engle's prediction of when short-term memory tasks should not depend on secondary memory. Indeed, there were no differences in performance between mismatch and no cue series when the series length was four (thus not exceeding the capacity of primary memory) and no arithmetic problems were given (thus not imposing any secondary task demands). This pattern is different from the one provided by the other three cells tested in the current investigation, and thus provides some evidence that the processes underlying performance in the simple-4 condition is qualitatively different from the complex/simple-8 and complex-4 conditions. In an interpretation consistent with the current aims of the study, it may be the case that simple-4 relies on primary memory, whereas the other three conditions rely on secondary memory. However, as will be discussed shortly, the pattern in the simple-4 condition still provides evidence for the encoding specificity principle, which potentially clouds the ability to use the encoding specificity principle to distinguish between contributions from primary and secondary memory. Further discussion of these results, in conjunction with the results from the first two experiments, will now be considered.

General Discussion

The goal of this investigation was to examine whether the secondary memory component involved in working memory tasks is the same as that used in long-term episodic memory tasks. In order to address this question, three experiments were conducted to determine whether recall on short term memory tasks is cue dependent, and if so, what (if any) are the boundary conditions for this phenomenon. Results from all three experiments indicate that recall in the mismatch condition was worse than in the no cue condition, consistent with the encoding specificity principle. Moreover, this pattern only emerged in conditions where performance would be expected to be dependent on secondary memory (e.g., Unsworth & Engle, 2007a, 2007c). Specifically, cue dependency was observed not only when secondary task demands were present on supraspan length series (Experiment 1), but also when no secondary task demands were present on supraspan length series (Experiment 2), as well as when secondary task demands were present on subspan length series (Experiment 3). Perhaps just as importantly, when no secondary task demands were present on subspan length series, a different pattern of recall was observed (mismatch = no cue). Taken together, these findings provide insight into Unsworth and Engle's (2007a, 2007c) dual-component model of working memory, which will now be discussed.

Implications for the Dual-Component Model of Working Memory

The results of this investigation are consistent with models of working memory that include contributions from secondary memory as a component as well as contributions from the focus of attention or primary memory (e.g., Baddeley, 2000; Cowan, 1999; Oberauer, 2002). More specifically, the present findings provide strong support for Unsworth and Engle's (2007a, 2007c) dual-component model of working memory, and evidence that the secondary memory component involved in working memory tasks is dependent on the same variables as long-term episodic memory tasks (e.g., Loaiza et al., 2011; Rose & Craik, 2012; Rose et al., 2014). The supporting evidence for the latter point comes from the fact that performance on a working memory task is cue dependent in ways similar to what is typically observed in studies of long-term episodic memory (e.g., Roediger & Adelson, 1980; Roediger & Payne, 1983; Thomson & Tulving, 1970). More specifically, when the retrieval cues are not the same as those given during encoding, recall is worse than when the retrieval cues are the same. Importantly, when the retrieval cues are not the same as those used during encoding, recall is worse than when no cues are provided at all. This pattern persists even when the retrieval cue is more strongly associated with the target word than with the encoding cue is (as was the case in the current investigation).

Although there is already evidence to suggest that manipulations affecting performance on long-term episodic memory tasks also affect performance on working memory tasks (e.g., Loaiza et al., 2011; Rose & Craik, 2012; Rose et al., 2014), the levels-of-processing manipulation used to support this position is encoding-based in nature. Results from the current set of experiments provide converging evidence that the secondary memory component involved is working memory and long-term episodic tasks are functionally the same using a manipulation that is also retrieval-based, though the encoding specificity principle also emphasizes the role of encoding conditions. This is particularly important, as encoding processes only constitute a subset of the approaches used to study memory, and assessing similarities between working memory and episodic memory tasks from both "ends" of the memory formation process provides much stronger explanatory powers.

The current investigation systematically examined when encoding specificity effects were observed as a function of task demands and series length, and in doing so, allowed testing of the predictions based on Unsworth and Engle's (2007a; 2007c) dual-component model of working memory. Specifically, Unsworth and Engle suggested that working memory represents the ability to retrieve information from secondary memory after it has been displaced from primary memory on immediate memory tests, and that performance on such tests relies on secondary memory when the attentional resources needed to maintain to-be-remembered items in primary are diverted to another task, or when the capacity of primary memory (four items) is exceeded.

However, it should be noted that these "rules" for when information is displaced into secondary memory have not received empirical attention with regards to assessing similarities between working memory tasks and long-term episodic memory tasks. That is, although studies have provided evidence that there are qualitative differences between information in primary and secondary memory by definition of Unsworth and Engle's "rules" (e.g., Unsworth & Engle, 2006), these studies have not provided evidence that these differences are because the same secondary memory component may be involved in performing both working memory and long-term episodic memory tasks. Conversely, studies that provide such evidence that the same secondary memory component may be involved in performing both tasks (e.g., Rose & Craik, 2012) have not demonstrated that this only occurs under the conditions specified by Unsworth and Engle's dual component model of working memory.

This distinction becomes theoretically important under the consideration that the two methods of displacement may not necessarily invoke similar demands, as may be implied by the Unsworth and Engle (2007a, 2007c) framework. Indeed, although performance on complex span tasks and longer trials of simple span tasks share some variability, they also account for independent variance in fluid intelligence (Unsworth & Engle, 2006). Furthermore, Loaiza and McCabe (2012) showed no benefit on a delayed memory test on longer series of a simple span task compared to shorter series, contrary to what is typically observed with complex span tasks (the "McCabe effect"; McCabe, 2008; see also Loaiza, Duperreault, Rhodes, & McCabe, 2014). This evidence for asymmetry between the two methods of displacement make it critical to assess similarities between working memory tasks and long-term episodic memory tasks as a function of the different methods of displacement described by Unsworth and Engle (2007a, 2007c).

Indeed, when either (or both) of these displacement conditions occurred in the context of the current investigation, performance in the mismatch condition was better than the no cue condition; when neither conditions occurred, performance between the two conditions did not differ. Put another way, encoding specificity (mismatch < no cue) on a short term memory task was only observed under situations where Unsworth and Engle predicted that secondary memory would be involved. Therefore, the key theoretical implication of the present findings is that secondary memory is involved in the performance of working memory tasks under the conditions specified by Unsworth and Engle, and that this secondary memory component may be the same as the one that underlies long-term episodic memory.

The dual-component model posits that items displaced from primary memory must be retrieved from secondary memory by way of a cue-dependent search process (e.g., Raaijmakers & Shiffrin, 1980). In situations where retrieval is required, Unsworth and Engle suggest that the ability to effectively delimit the search process will influence retrieval success, and that individuals with higher working memory ability are better at limiting the search set to only relevant information through the use of different cues. Further, Unsworth and Spillers (2010) have suggested that individuals with higher working memory ability are more likely to engage in strategic encoding processes. Thus, it may be the case that the efficiency of working memory reflects the extent to which individuals can reinstate the original encoding cues during retrieval. If this is the case, it would be expected that if the cues presented during recall on a working memory task do not match those during encoding, performance would suffer.

This view is consistent with findings by Unsworth, Brewer, and Spillers (2011), who reported that individuals who performed better on a working memory task were hurt more by mismatch cues during recall in an encoding specificity paradigm than individuals who performed more poorly on the working memory task. Along with the results of this investigation, results by Unsworth et al. suggest that the ability to use cues to encode to-be-remembered information and facilitate recall at a later time is important in a working memory task, and that retrieval from secondary memory on working memory tasks is fraught with the same potential problems (e.g., proactive interference, encoding deficiencies, output interference) as retrieval from secondary memory on episodic long-term memory tasks.

As predicted, performance in three tasks of this investigation (complex/length-8, simple/length-8, and complex/length-4) yielded poorer performance with mismatch cues than both match cues and when there were no cues, and this was taken as evidence for encoding specificity. It is important to note that the one task that yielded no difference between the mismatch and no cue conditions (simple/length-4) is not inconsistent with a strict interpretation of the encoding specificity principle put forth by Thomson and Tulving (1970), who stated that recall under mismatch conditions should not be higher than under no cue conditions (p. 258). However, given the context of their study, it is fair to say that their interpretation applies to retrieval from secondary memory. In the simple/length-4 condition, the information being recalled is most likely retrieved from primary memory, and not secondary memory. If that is the case, given evidence that information stored in primary memory is not encoded on a semantic dimension (e.g., Baddeley & Dale, 1966), then being presented with a semantically associated cue that was not previously seen during encoding would not be any more or less helpful for recall than being provided with no cues at all. Thus, one would expect that recall would not differ between the mismatch and no cue conditions. Moreover, this reasoning is consistent with the present claim that the encoding specificity effect (when defined as the comparison between mismatch and no cue conditions) is stronger under conditions where recall relies on secondary memory than conditions where recall relies only on primary memory.

That said, one can make an argument that the encoding specificity principle does in fact hold in primary memory when comparing performance on match and mismatch series for participants in the simple span condition. Recall that although the match condition was dropped in Experiment 3, the practice series were still done using match cues. Furthermore, Experiments 1 and 2 demonstrated that performance on the practice series was not different from performance on the test series for the match condition. As a result, we can compare performance on the practice series (match) with performance on the mismatch test series to further assess the encoding specificity hypothesis. Indeed, when doing so, performance on the match series (M = .85) was greater than the mismatch series (M = .55), $t(49) = 6.65$, $p < .001$, d = 1.37. It should also be noted that if a follow-up study was conducted where strong cues served as match cues and weak cues served as the mismatch cues, we would likely observe the same pattern, but with even larger effect sizes (Thomson & Tulving, 1970). This hypothetical interaction between encoding and retrieval cues can be taken as support for the encoding specificity hypotheses, and in the context of the current investigation, may suggest encoding specificity is in fact observed in primary memory, at least when defined as the comparison between match and mismatch conditions. What is less debatable, however, is that the processes underlying performance on simple/length-4 series is qualitatively different from the other three series types (complex/length-8, simple/length-8, and complex/length-4).

In contrast, it is worth pointing out the difference in recall between the mismatch and no cue condition did not differ much between the various instances when recall was thought to rely on secondary memory. That is, secondary task demands (Experiment 3; Cohen's d = .50) did not produce a greater effect size than a supraspan series length (Experiment 2; d = .58), and a similar effect size was observed when secondary task demands and a supraspan series length were combined (Experiment 1; d = .60). Indeed, a 2 (mismatch vs no cue) x 3 (Experiment 1 vs Experiment 2 vs Experiment 3) ANOVA confirmed that the difference in recall between mismatch and no cue condition did not differ across the experiments, $F < 1.0$. One way to interpret this is that there is a threshold such that once it is crossed, performance on a working memory task relies primarily on secondary memory. Moreover, the similar effect sizes observed with supraspan series lengths (Experiment 2) and secondary task requirements (Experiment 3) suggest that the qualitative nature of the displacement to secondary memory is the same, regardless of what causes it (cf. Loaiza & McCabe, 2012).

Relation to Higher-Order Cognitive Ability

It is important to note that while the interpretation of the results of this investigation is predicated on the assumption that the modified span task used here does in fact assess working memory, the use of word-pairs in this task raises the question of whether it actually measures working memory or whether it measures another construct. One approach to answering this question is to examine the relationship between performance on this task and performance on other known measures of working memory, as well as between performance on the modified span task and other higher-order constructs known to be related to working memory (e.g., fluid intelligence). Fortunately, this approach is feasible here because a pilot sample (n = 60), who were studied in a laboratory setting in the early stages of this investigation, was tested on a working memory task (operation span task) and a fluid intelligence task (Raven's Progressive Matrices; Raven, 1936), as well on the version of the present modified span task that included a secondary task (Experiment 1).

If performance on this modified complex span task does reflect working memory, then it should be positively correlated not only with the operation span task, but also with the fluid intelligence task (e.g., Engle et al., 1999; Kane & Engle, 2002). Indeed, results from the pilot sample were consistent with what would be predicted under the assumption that the modified span task provides a valid assessment of working memory: Performance on the series from the match condition of the modified span task (where the cue at retrieval matches the cue at encoding) was strongly correlated with performance on the operation span task ($r = .58$, $p < .001$), and was also significantly correlated with performance on the Raven's Progressive Matrices task ($r = .34$, $p = .008$). The observed pattern of correlations indicates that the use of word-pairs as to-be-remembered items did not change the nature of the span task.

Next, it cannot be known for certain that the word-pairs in the current set of experiments were treated as two different to-be-remembered items. Indeed, typical complex span tasks present only one to-be-remembered item at a time, making it unambiguous as to how many items a participant is encoding from any given series. Depending on the series length in these experiments, it is assumed that participants encoded either eight or four distinct items, though they were technically shown sixteen and eight distinct words, respectively. However, there is extensive research to suggest that individuals will attempt to "chunk" several pieces of information together into larger units in order to help them remember the information better (e.g., Miller, 1956). For example, trying to remember a series of numbers like "*1-0-2-7-2-0-1-1-6-1-0-9-1-1*" that exceeds memory span capacity is easier for most St. Louis Cardinal fans if they recognize that the series can be broken down into the date "*10-27-2011*" that the Cardinals beat the Rangers in game "*6*" of the World Series by a score of "*10-9*" in the "11^{th}" inning. With regards to the present investigation, because the word-pairs were semantically related to one another (albeit weakly), it seems likely that participants were encoding each word-pair as a single chunk of information. That said, future studies may consider using strongly related word-pairs to better ensure that participants are treating each word-pair as a single unit of information (i.e., chunking).

The results of this investigation provide converging evidence that working memory tasks rely on the same secondary memory component as long-term episodic memory tasks. This finding raises the question of why, if both types of task primarily assess secondary memory, researchers should not simply use the latter type of task, as episodic memory tasks presumably provide a better measure of retrieval processes than do working memory tasks. This was the question posed by Mogel, Lovett, Stawski, and Sliwinski (2008), who gave participants a battery of tasks that measured working memory, primary memory, secondary memory (using long-term episodic memory tasks), and fluid intelligence. Structural equation modeling revealed that after accounting for variance due to working memory, secondary memory still explained a significant amount of unique variance in fluid intelligence. In contrast, after accounting for the variance from secondary memory, working memory did not explain any unique variability in fluid intelligence. Mogel et al. took these results to mean that the well-known relationship between working memory and fluid intelligence is not as “special” as previously thought, as it is driven primarily by differences in episodic memory.

However, a subsequent study by Shelton, Elliot, Matthews, Hill, and Gouvier (2010) provided contrary evidence suggesting that working memory does explain unique variance in fluid intelligence after all, and that it captures all the variance accounted for by secondary memory. Shelton et al. proposed that the differences between their findings and those of Mogel et al. (2008) may have been due to task selection (recall vs. recognition tasks for secondary memory) or how the working memory tasks were administered (experimenter- vs self-paced). Regardless of the source of the discrepancies, the Shelton et al. findings suggest that, at least with the tasks that they used, secondary memory is only be partly responsible for the observed link between working memory and fluid intelligence, and that working memory may in fact be “special”. Furthermore, the evidence from Shelton et al. (2010) converges with the results of Unsworth, Brewer, and Spillers (2009), who also found that working memory accounted for unique variance in fluid intelligence after controlling for individual differences in secondary memory ability. With regard to the current investigation, although the present results provide converging evidence that working memory relies on the same secondary memory component as long-term episodic memory, it is a separate question whether retrieval from secondary memory constitutes the full scope of what working memory tasks measure, or merely a subset of it.

Alternative Explanations

Another question to consider is how often participants in the mismatch and no cue conditions recalled the original cue words that were presented along with the target words (Thomson & Tulving, 1970). In considering this question, however, one needs to bear in mind that the amount of information encoded (and thus the number of likely response alternatives) may differ between the three cue conditions utilized in the three reported experiments. In the match condition, for example, there was only one type of item from the preceding series (i.e., the target words) that were likely responses at recall. In the mismatch and no cue conditions, in contrast, where the participant was presented with either new (i.e., mismatched) cues or no cues at recall, there were two types of items that had been presented as part of the series (i.e., the original cues and target words) that then were likely responses at recall. It is conceivable that this difference in the number of possible responses may in part explain why recall in the match condition was better than in the other two conditions, though it would not seem to account for why recall in the mismatch condition was worse than the no cue condition.

Collapsing across the experiments in which encoding specificity was observed (Experiments 1, 2, and the complex span condition in Experiment 3), the proportion of study cues recalled in the mismatch condition (*M* = .20, *SD* = .21) was significantly greater than in the no cue condition (*M* = .15, *SD* = .19), $t(248) = 2.15$, $p < .05$, $d = 0.27$. If participants in the mismatch condition were more likely than those in the no cue condition to recall the original cue word (as this analysis confirms), one might expect that they would be more likely to recall the target word, since they are reinstating the cue used during encoding. However, recalling the study cue does not seem to have had such facilitating effects on recall of target words.

Another question to consider is why performance for the no cue condition in the simple-4 series was so low. Indeed, the results from Experiment 3 indicate that when participants were given four words to remember, they can only recall about two words (*M* = .52) when prompted immediately after presentation. This is surprising, because it suggests such a low limit to participants' short-term memory capacity, and stands in stark contrast to the near perfect performance in similar conditions reported by Unsworth and Engle (2006). There are two possible explanations for this. The first is that recall procedures used in these tasks were less than ideal for assessing memory. More specifically, participants were asked to remember word-pairs, yet during recall, were only asked to recall the target word. It may be the case that recalling the target word is made harder by the fact that they were encoded together with the cue word, compared to having to recall the target word if no cue was present at study. Indeed, this is consistent with results found by Thomson (1972), who found that target words were better recognized if they had been presented alone during study compared to if they were presented along with a cue.

The second possible explanation for why performance in the no cue condition in the simple-4 series was so low was because the initial practice series habituated participants into looking for cues during recall, such that they were hurt by the lack of available cues during recall on the test series. Put another way, the low performance may be attributed to the fact that participants were expecting cues during recall in the test series. This explanation does have some empirical support from another sample ($N = 15$) where participants were given simple-4 series with no practice series. Under these conditions, recall in the no cue condition was substantially higher ($M = .73$) than when practice series were administered. Taken together, there are two plausible explanations for the low performance in no cue condition in the simple-4 series that future studies may be able to confirm.

Turning to the final question, there is the question of whether participants were using the provided cues to help remember the target words. To the extent that participants were making the associative connections between the cue and target, recall on mismatch series should have suffered. As such, it was important that recall in the first four series was done using match cues, in order to instill confidence in the participants that they could form associations to help aid recall. Indeed, when such practice trials have not been used, no differences are seen between mismatch and no cue conditions (Thomson & Tulving, 1970, Experiment 1). One way of checking whether this was done is to examine whether the change in performance across the multiple practice series predicts performance on the final test series for the mismatch condition. It would not be surprising to see participants do better on the last series than the first series, if only because of practice effects. However, participants could have also been improving because they began to rely more on the cues to help learn the target words during the encoding phase. Put differently, participants on the final practice series may have been more likely to use make semantic associations between the cue and target than on the first practice series. If this is the case, participants who relied more on making these semantic associations across the four practice series should have been hurt most when the cue on the test series was changed (i.e., in the mismatch condition).

This was confirmed by aggregating data from all the participants in the experiments that showed an encoding specificity effect, and excluding participants who scored higher than .88 on the first series. This was done because the excluded participants were so close to ceiling that it would have been difficult to see changes in their performance across the practice series, which would potentially have weakened any observed patterns. Recall scores were standardized within each experiment, after which the correlation between change scores (recall on the fourth series minus recall on the first series) and recall on the final series was examined. Analyses revealed a correlation between change scores and performance on the test series, $r(112) = -.21$, $p = .027$. In contrast, the corresponding correlations in the match condition ($p = .188$) and no cue condition ($p = .335$) were not significant. The negative correlation observed in the mismatch condition confirms that the participants who were doing more poorly on mismatch series were those who were also improving the most across the practice series. Moreover, the lack of such correlations in the other two conditions suggests that this improvement across practice series was likely due to some sort of increased reliance on binding the cue and target, as expected in this investigation.

Closing Thoughts

Unsworth and Engle (2007a, 2007c) proposed that when one performs a working memory task, information brought back to primary memory from secondary memory is done so via a cue-dependent search process. But in order for to-be-remembered information to require retrieval from secondary memory, it first has to be displaced from primary memory. Under these circumstances, the present results suggest that retrieval from secondary memory during performance of a working memory task is cue-dependent. Furthermore, the present results demonstrate that this only occurs under conditions where Unsworth and Engle (2007a; 2007c) predict that to-be-remembered information will be displaced from primary memory. Thus, this investigation sheds light on the qualitative nature of the retrieval processes that take place in working memory tasks, and the extent to which it is similar to the way we normally think about retrieval on long-term episodic memory tasks. In short, modifying a standard working memory task allows for the assessment of the degree to which performance on working memory tasks "behaves" like performance on long-term episodic memory tasks, thereby contributing to a better understanding of what working memory and its components truly are.

Appendix A.

TARGET	WEAK CUE	STRONG CUE
SUM	product	total
JACKET	shirt	vest
TABLE	lamp	chair
HOLE	plug	crater
OYSTER	fish	clam
EMPTY	vague	fill
NOISE	crowd	loud
HEN	goose	rooster
PIE	dessert	custard
LEATHER	boot	suede
MOVIE	comedy	preview
HOLD	pause	grip
VOTE	candidate	election
POT	bowl	marijuana
LADDER	hook	rung
BROTHER	family	sibling
BOSS	leader	manager
SYRUP	sticky	waffles
POLICE	legal	badge
SQUARE	geometry	cube
LAND	mass	property
BAR	club	pub
TIME	life	era
ROD	steel	reel
SILVER	tray	tarnish
SANDWICH	chicken	bologna
REEL	movie	rod
BITE	apple	snake
BLOOD	hound	dracula
OVEN	roast	microwave
TRACK	road	railroad
SISTER	mother	sibling
METAL	wood	steel
RAT	cheese	mice
FLEA	market	tick
COLOR	white	race
FIELD	battle	plow
TILE	carpet	heading
SODA	fountain	beverage
CLUB	country	member

Appendix B.

Instructions before the start of the task

On each trial of this task, you will see a series of word-pairs that you must remember. For example, you may given the word-pair 'market : FLEA'. In this instance, 'market' is the cue word, and 'FLEA' is the target word that you will be asked to remember at the end of each trial. In addition, after the presentation of each word-pair, you will see a math equation that you must read and decide whether it is correct, indicating your decision by pressing the appropriate button. Please make your decision as quickly as possible without sacrificing accuracy.

At the end of each trial, the top of the screen will display the word 'RECALL' in red. This is your indication to recall the target words ('FLEA') in the text box, in the exact order in which they were presented.

You are to type in only one word in the text box. Once you are done typing that word, hit Enter, and a new blank text box will appear on the screen.

Most importantly, for recall of each target word, the blank text box will be accompanied with a cue word related to the target item you need to recall. For example, the word 'market' may appear next to the blank text box. In this case, you would type in the target word 'FLEA'. Keep in mind that although the cue presented at recall will almost always match the cue seen earlier during the trial, this may not always be the case. However, the cues will always be related to the target word that belongs in the text box. Please do your best to recall the target word, although if you cannot remember, you may type in your best guess. Keep in mind that, during recall, only recall the word that was on the right side (the capitalized target word). In other words, you will never be asked to recall the (lowercase) cue word on the left.

Please do not use any external aid to write down the presented word-pairs. This task is designed to assess certain aspects of human memory, and your performance will in no way change the length/difficulty of this task, nor will it affect your compensation.

Instructions during recall for the mismatch condition

Although these are different cues, they are semantically related to the target word you are supposed to recall (the capitalized word on the right).

Instructions during recall for the no cue condition

No cues are provided for this series. Please recall all of the capitalized words (that were presented on the right) in any order.

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, *2*, 89-195.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal of Experimental Psychology*, *18*, 362-365.

Baddeley, A. D. (1972). Retrieval rules and semantic coding in short-term memory. *Psychological Bulletin*, *78*, 379-385.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417-423.

Baddeley, A. D., & Dale, H. C. A. (1966). The effect of semantic similarity on retroactive interference in long-and short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *5*, 417-420.

Baddeley, A. D., & Hitch. G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior,14*, 575-589.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior, 13,* 471-481.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173-189.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology, 26,* 467-482.

Bui, D. C., Friedman, M. C., McDonough, I. M., & Castel, A. D. (2013). False memory and importance: Can we prioritize without consequence? *Memory & Cognition, 41,* 1012-1020.

Bui, D.C., Maddox, G., & Balota, D. (2013). The roles of working memory and intervening task difficulty in determining the benefits of repetition. *Psychonomic Bulletin & Review, 20(2),* 341-347.

Bui, D. C., Maddox, G. B., Zou, F., & Hale, S. (2014). Examining the lag effect under incidental encoding: Contributions of semantic priming and reminding. *Quarterly Journal of Experimental Psychology.*

Bui, D. C., Myerson, J., & Hale, S. (under review). *Demonstrating age-related declines in processing speed using online samples.*

Bui, D. C., & Myerson, J. (2014). The role of working memory abilities in lecture note-taking. *Journal of Learning and Individual Differences, 33,* 12-22.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15,* 17-31.

Cowan, N. (1999) An embedded-processes model of working memory. In: Miyake A. and Shah P. (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive*

Control. Cambridge University Press, Cambridge, UK, pp. 62–101.

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

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology: general, 104*, 268.

Crowder, R. G. (1976). *Principles of Learning and Memory.* Hillsdale, N.J.: Erlbaum.

Daneman, M., & Carpenter, P. A., (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3,* 422-433.

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

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309-331.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1-16.

Goldman-Rakic, P. S. (1992). Working memory and the mind. *Scientific American, 267,* 110–117.

Haberlandt, K. (1997). *Cognitive psychology* (2nd ed.). Boston: Allyn & Bacon.

Jonides, J., Lewis, R., Nee, D., Lustig, C., Berman, M., & Moore, K. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, 59, 193–224.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic bulletin & review, 9*, 637-671.

Kintsch, W., & Buschke, H. (1969). Homophones and synonyms in short-term memory. *Journal of Experimental Psychology, 80*, 403-407.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence, 14*, 389-433.

Loaiza, V. M., Duperreault, K. A., Rhodes, M. G., & McCabe, D. P. (2014). Long-term semantic representations moderate the effect of attentional refreshing on episodic memory. *Psychonomic Bulletin & Review*.

Loaiza, V. M., & McCabe, D. P. (2012). Temporal contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, *40*, 191-203.

Loaiza, V., McCabe, D., Youngblood, J., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 5, 1258–1263.

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language, 58,* 480-494.

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*, 222-243.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63,* 81–97.

Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *The Journal of General Psychology*, *50*, 129-139.

Miyake, A., & Shah, P. (Eds.) (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological Science*, *19*, 1071-1077.

Multhaup, K. S., Balota, D. A., & Cowan, N. (1996). Implications of aging, lexicality, and item length for the mechanisms underlying memory span. *Psychonomic Bulletin & Review, 3*, 112-120.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Newman, S. E., & Frith, U. (1977). Encoding specificity vs associative continuity. *Bulletin of the Psychonomic Society*, *10*, 73-75.

Nilsson, L. G., Law, J., & Tulving, E. (1988). Recognition failure of recallable unique names: Evidence for an empirical law of memory and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 266-277.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological review*, *88*, 93-134.

Raven, J. C. (1936). *Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive.* MSc Thesis, University of London.

Roediger, H. L., & Adelson, B. (1980). Semantic specificity in cued recall.*Memory & Cognition*, *8*, 65-74.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*, 803-814.

Roediger, H. L., & Payne, D. G. (1983). Superiority of free recall to cued recall with “strong” cues. *Psychological research*, *45*, 275-286.

Rose, N. S., Buchsbaum, B. R., & Craik, F. I. (2014). Short-term retention of a single word relies on retrieval from long-term memory when both rehearsal and refreshing are disrupted. *Memory & Cognition*, *42,* 689-700.

Rose, N. S. & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from a levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 38, 4, 1019-1029.

Rose, N. S., Myerson, J., Roediger, H. L., & Hale, S. (2010). Similarities and differences between working memory and long-term memory: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 2, 471-483.

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review, 15,* 364-371.

Shallice, T. & Warrington, E.K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *Quarterly Journal of Experimental Psychology*, *22*, 261–273.

Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B. D., & Gouvier, W. D. (2010). The relationships of working memory, secondary memory, and general fluid intelligence:

working memory is special. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 813-820.

Shulman, H. G. (1970). Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 499-508.

Shute, V. J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research, 7,* 1-24.

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language*,*40*, 83-108.

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 20,* 261-288.

Tamez, E., Myerson, J., & Hale, S. (2012). Contributions of associative learning to age and individual differences in fluid intelligence. *Intelligence, 40,* 518-529.

Thomson, D. M. (1972). Context effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior, 11,* 497-511.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology, 86,* 255–262.

Tulving, E. (1983). *Elements of Episodic Memory*, New York: Oxford University Press.

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

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381-391.

Tulving, E. (1974). Cue-Dependent Forgetting: When we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *American Scientist*, 74-82.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127-154.

Unsworth N., Brewer G.A., Spillers G.J. (2009). There's more to the working memory capacity-fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review,16*, 931–937.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2011).Variation in working memory capacity and episodic memory: Examining the importance of encoding specificity. *Psychonomic Bulletin & Review, 18*, 1113-1118.

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*, 68-80.

Unsworth, N., & Engle, R. W. (2007a). Individual differences in working memory capacity and retrieval: A cue-dependent search approach. *The foundations of remembering: Essays in honor of Henry L. Roediger, III*, 241-258.

Unsworth, N., & Engle, R. W. (2007c). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological review*, *114*, 104-132

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17*, 200-205.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory short term memory. *Brain*, *92*, 885–896.

Watkins, M. J. (1974). Concept and measurement of primary memory. *Psychological Bulletin*, *81*, 695-711.

Watkins, M. J. (1977). The intricacy of memory span. *Memory & Cognition*, *5*, 529-534.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 442-452.

Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior*, *13*, 430-447.

Waugh, N., & Norman, D. (1965). Primary memory. *Psychological Review, 72*, 2, 89-104.