Rehearsal Processes

Geoff Ward, Department of Psychology, University of Essex, UK.

Abstract
This work reviews rehearsal processes in human memory. It considers the different types and functions of rehearsal, the different methods used to examine rehearsal, and considers the role of rehearsal in classic short-term and working memory tasks and theories, including the Brown-Peterson task, immediate free recall, immediate serial recall, and the complex span task. A variety of different types of rehearsal have been proposed, including: maintenance (or articulatory) rehearsal, elaborative rehearsal, attentional refreshing, covert retrieval, and short-term consolidation. Theorizing about rehearsal can be highly contentious and the chapter discusses theories that assume rehearsal enhances later accessibility, theories that assume rehearsal merely maintains items in a highly accessible state, and theories that propose that there is no causal link between rehearsal and recall. It considers the relationship between rehearsals and later recall, the relationship between rehearsals and repetitions, and considers whether rehearsal and recall are underpinned by the same retrieval mechanisms. It concludes with a summary of the current points of contention and a personal viewpoint: that rehearsal is most likely to enhance recall to the extent that retrieval enhances later retrieval, because rehearsal and recall are underpinned by the very same mechanisms.

Keywords: rehearsal, elaboration, repetition, retrieval, free recall, serial recall, complex span, short-term memory, Working Memory, attentional refreshing,

Introduction: what is rehearsal?

When participants are asked to recall a short sequence of items for a later test, they often repeat earlier items to themselves during the presentation of later items, a process known as *rehearsal*. As Johnson (1980) argues, rehearsal has been assigned a major role in practically all memory methodologies, and in this chapter, I consider the different proposed functions of rehearsal, the different methods that have been used to examine rehearsal, and the different types of rehearsal that have been proposed. I will concentrate on the proposed role of rehearsal in many classic theories of memory tasks and phenomena. A theorist's understanding of rehearsal is affected by their preferred account of how repetitions are represented in memory, their preferred understanding of working memory capacity and serial position effects, and by the theorist's preferred account of forgetting. It is therefore unsurprising that theorising about rehearsal is contentious, and it is understandable why accounts of rehearsal are best understood within theories of particular tasks and findings.

In this chapter, I will first consider the proposed functions of rehearsal, the different methods for studying rehearsal, and the different types of rehearsal. I will then discuss the proposed roles of rehearsal in four influential short-term or working memory tasks: the Brown-Peterson task, free recall, serial recall, and the complex span task. I will move on to discuss the role of rehearsal in long-term episodic memory: the relationship between rehearsal and long-term retention and the relationship between rehearsals and repetitions. I will end the review with a summary of the current points of contention in the rehearsal literature and a personal perspective regarding rehearsal.

Proposed functions of rehearsal

The proposed functions of rehearsal differ from task to task, but rehearsal is generally considered to be under voluntary control and to offer mnemonic advantages for those items that are rehearsed. As we will see, rehearsal has been argued (i) to increase the strength of associations within long-term memory (Atkinson & Shiffrin, 1968, 1971; Raaijmakers & Shiffrin, 1981; Unsworth & Engle, 2007), (ii) to maintain items in serial order in a highly accessible state at close to their initial level of activation by offsetting forgetting due to trace decay (e.g., Baddeley, 1986; Barrouillet, Bernandin & Camos, 2004; Oberauer & Lewandowsky, 2011; Reitman, 1974; Towse, Hitch & Hutton, 2000; Unsworth & Engle, 2007), and (iii) when rehearsal is prevented, it is argued that we see the full effects of short-term forgetting (Brown, 1958; Peterson & Peterson, 1959).

These mnemonic advantages may arise in multiple ways. Rehearsal offers the opportunity to strengthen item-item associations (Raaijmakers & Shiffrin, 1981), rehearsal of subsequences may help promote temporal groupings (Ryan, 1969; Wickelgren, 1964, 1967), and the presentation of one item might lead to the rehearsal of related earlier items, encouraging semantic or associative clustering (Rundus, 1971). Items that are rehearsed may create multiple copies, may be co-rehearsed with different sets of items, and may be re-ordered and distributed throughout the list including towards the end of the list (Tan & Ward, 2000). Such rehearsal schedules may increase the amount of distributed practice that an item experiences (Modigliani & Hedges, 1987), and will reduce the functional retention interval of that item (Brodie & Murdock, 1977).

The act of rehearsal can also change the representation of what is encoded. When visual words are read silently, then rehearsal can lead to the phonological recoding of pronounceable visual stimuli (Baddeley, 1966; Estes, 1973). Rehearsal of unfamiliar items can encourage long-term acquisition of the stimuli (Baddeley, Gathercole, & Papagno, 1998), and in children this has been argued to lead to vocabulary acquisition (Gathercole & Baddeley, 1989). Although rehearsal is most commonly associated with the post-stimulus repetition and processing of verbal stimuli, many theorists argue for accounts of rehearsal to also include non-verbal stimuli such as visual (Watkins, Peynircioglu, & Brems, 1984) and visuo-spatial stimuli (e.g., Awh & Jonides, 2001; Awh, Jonides & Reuter-Lorenz, 1998; Cortis, Dent, Kennett, & Ward, 2015; Cortis Mack, Dent & Ward, 2018).

Methods for studying rehearsal

In a typical experiment, there may be multiple opportunities to rehearse: rehearsal of an item can take place immediately after its presentation, during the inter-item intervals between later items, or during the retention interval prior to test. Unfortunately, since rehearsal may be covert in many studies, one can rarely say with any certainty whether or not rehearsal was actually performed by any given individual on any trial. Nevertheless, a number of experimental manipulations have been commonly used to vary the opportunity for rehearsal. We will consider these methods in more detail as we look at specific tasks in later sections. However, since each type of manipulation has advantages and disadvantages, an overarching point is that it is wise to study rehearsal using a range of different methods and techniques.

The first type of manipulation is at the level of the to-be-remembered item; one can manipulate rehearsal if one assumes that some stimuli are easier to rehearse than others. For example, long, multisyllabic words take longer to say and so are less efficiently rehearsed than shorter, monosyllabic words (Baddeley, Thomson & Buchanan, 1975). Similarly, some stimuli are harder to pronounce than others (e.g., Meunier, Stanners & Meunier, 1971). Although there are clear recall advantages between many sets of easier-to-rehearse and harder-to-rehearse stimuli, a difficulty with this method is that it is hard to be certain whether the mnemonic difference is entirely attributable to the difference in the ease of rehearsal.

A second type of manipulation is to vary the presentation schedules. A slow presentation rate affords greater opportunities to rehearse than a faster presentation schedule (e.g., Tan & Ward, 2000, 2008). One complication is that varying the presentation rate will also vary the retention interval for unrehearsed items.

A third type of manipulation involves changes to the concurrent processing requirements that reduce the opportunities to rehearse. For example, participants may be required to repeatedly utter irrelevant verbal items such as “the the the” or “ba ba ba” during the presentation of list items, a procedure that is commonly termed concurrent articulation or articulatory suppression (Murray, 1968). Other manipulations include requiring participants to respond to a concurrent processing task using oral responses rather than by button presses (e.g., Camos, Lagner & Barrouillet, 2009). However, a complication with this manipulation is that it is uncertain whether the requirement to repeatedly utter irrelevant speech does more than simply restrict rehearsal (Lewandowsky & Oberauer, 2015).

A fourth type of manipulation is to vary task instructions. For example, participants who are uninformed about a later memory test and who are processing individual items incidentally are less likely to rehearse than those who are fully anticipating a later test and so who process the items intentionally. One complication is that it is often difficult to be certain that the allocation of goal-directed attention during item presentation is matched between incidental and intentional learning conditions.

In all the manipulations so far presented, it is assumed that the degree of rehearsal varies, but one cannot be certain because the rehearsals themselves are not directly observed. A more direct approach is to instruct participants to rehearse aloud, a procedure known as the overt rehearsal technique (Rundus & Atkinson, 1970). Using this technique, the experimenter can gain better understanding of the patterns of rehearsal on any trial and so can compare the differences in rehearsal with the differences in recall at the level of groups, individuals, and even trials. One concern with this method is that the overt rehearsal instructions may encourage participants to encode the stimuli differently from how they might otherwise have done, such that the overt rehearsal method provides good data on the overt rehearsal version of the task, but this may not be the same as the more standard version of the task when participants are free to encode the list in whatever way they wish. A second concern with the overt rehearsal method is that any observed relationship between patterns of rehearsals and patterns of recalls is correlational in nature rather than causal (Lewandowsky & Oberauer, 2015; Souza & Oberauer, 2018, 2020): a first list of words may be rehearsed more and recalled better than a second list of words because the material in the first list was easier to learn in some way or because a participant paid more attention during the first list than the second. Similarly, one participant may rehearse more and recall more than a second participant because they have superior memory or superior abilities in goal-directed attention. In each case, the cause of the positive relationship between rehearsal and recall lies in some additional factor that affects both measures.

A causal explanation can be established by experimentally manipulating the patterns of ‘rehearsals’ that are presented to participants. One approach is to instruct participants how to rehearse. For example, participants who are given *fixed* rehearsal instructions are required to rehearse only the current or most recently-presented item (e.g., Fishler, Rundus, & Atkinson, 1970; Glanzer & Meinzer, 1967), whereas those participants who are given *cumulative rehearsal* instructions are required to rehearse after each successive list item all the items that had been presented to date in forward order (e.g., Souza & Oberauer, 2018).

A causal link between rehearsal and recall can also be determined by presenting a participant in one group with the rehearsals generated by a yoked participant in a second group. In some circumstances, the participant is presented with the original list of words and its rehearsals; in other circumstances, the rehearsal schedule of one participant is applied to a different set of stimuli. In this way, the effects of rehearsal schedule can be separated from the memorability of the individual items themselves. A related method is to present participants with patterns of ‘rehearsals’ that have been generated by the experimenter. Experimenter-generated schedules allow the experimenter to exert more control, allowing researchers to experimentally manipulate the frequency, recency and distribution of rehearsals (e.g., Tan & Ward, 2000). The experimenter may alternatively generate the rehearsals using an algorithm, if one wanted to test putative models of the rehearsal process. Although presenting participants with experimenter-generated schedules has many benefits, a concern is whether an experimenter-generated presentation of a ‘rehearsal’ (i.e., a repetition) of a stimulus is always equivalent to a participant-generated rehearsal.

Finally, one can examine the effects of different encoding manipulations such as rehearsal by conducting post-task interviews in which the experimenter asks participants about the strategies that they had used. The participants' recall can then be categorised by their chosen strategy and the effects of strategy use on recall can be compared. One advantage of this technique is that it can determine the range of strategies that are spontaneously used, but the technique is only as effective as the accuracy of the participants' introspections and their memory of strategy use when assayed during the post-task interviews.

Types of Rehearsal

Before considering the proposed roles of rehearsal in different tasks and phenomena, it is important to acknowledge that most researchers assume that there are different types of processing that can occur during rehearsal. For example, Rundus (1971) argued that the overt rehearsal method that was discussed in the previous section was a method by which participants say out loud whatever earlier items came to mind during the presentation of the list. He argued that these rehearsals could result from a variety of sources, including rote rehearsal, visual imagery, and different types of mnemonics. Although acknowledging that the rehearsals had heterogeneous sources, he was nevertheless more concerned with relating the patterns of rehearsal with the patterns of recall, and less concerned with classifying each rehearsal by its source or origin.

By contrast, Craik and Lockhart distinguished between two forms of rehearsal: *maintenance rehearsal* (rote repetition) and *elaborative rehearsal* (repetition that involves the integration of the stimulus or its enrichment by association with other pre-existing memories, e.g., Craik & Tulving, 1975). Craik and Lockhart argued that memory is a by-product of processing, and that items that were processed in a deep way (i.e., using processing tasks that encouraged greater processing for meaning) would be remembered better than those that were encoded in a shallow way (i.e., using processing tasks that involved superficial or perceptual processing). Maintenance rehearsal was argued to retain the item at its original level of processing, whereas elaborative rehearsal would result in deeper processing leading to an improvement in later memory performance.

Developments in working memory have further highlighted the need to clarify the different processes that can be performed during rehearsal (for a recent review, see Oberauer, 2019). A distinction has been made between rote or *articulatory rehearsal* and *elaborative rehearsal* (e.g., Bartsch, Singmann & Oberauer, 2018; Lewandowsky & Oberauer, 2015), but it is also apparent that some form of working memory maintenance is possible even when verbal rehearsal is prevented, and the term *attentional refreshing* is commonly used to describe a domain-general mechanism capable of reviving memory traces by simply attending to them (Camos, Johnson, Loaiza, Portray, Souza, & Vergauwe, 2018; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Vergauwe & Langerock, 2017). In addition, there is increasing evidence for the role of retrieval from long-term memory in working memory tasks when rehearsal and refreshing is disrupted (e.g., Rose, Buchsbaum, & Craik, 2014). Finally, the term *short-term consolidation* or *working memory consolidation* has been used to refer to the additional post-stimulus processing that occurs after the offset of a brief stimulus (Bayliss, Bogdanovs & Jarrold, 2015; de Schrijver & Barrouillet, 2017; Ricker & Cowan, 2014; Ricker, Nieuwenstein, Bayliss, & Barrouillet, 2018).

The Role of Rehearsal in Short-term Memory and Working Memory Tasks

It is difficult to understand fully the proposed functions of rehearsal within different theories of memory without considering the tasks and findings that these theories seek to explain. The four main tasks that are focussed upon here are: (i) the Brown-Peterson task, (ii) the free recall task, (iii) the immediate serial recall task, and (iv) the complex span task. It is not a coincidence that all four tasks have been highly influential in developing the concepts of short-term and working memory: it is often assumed that rehearsal offsets the effects of trace decay that would otherwise cause short-term forgetting. However, as we proceed through our tour of classic memory tasks and theories, we will encounter additional and alternative proposed functions of rehearsal and we will detail the variety of different types of rehearsal processes that have been proposed.

Rehearsal in the Brown-Peterson task

In an early review of rehearsal and repetition, Bjork's (1970) primary interest concerned data relating to the Brown-Peterson task, a method used to examine short-term forgetting. In a typical trial using the Brown-Peterson task (Brown, 1958; Peterson & Peterson, 1959), participants are presented with a short sequence of verbal items, such as the consonant trigram "XJN" and they must try to recall these three letters in the correct serial order, following a delay of variable length. If an unfilled delay is used, and participants are free to rehearse during the retention interval, then a short sequence can be well remembered (e.g., Brown, 1958; Meunier, Ritz, & Meunier, 1972). However, if rehearsal is prevented by a backwards-counting task (e.g., continually subtracting 3 out loud from a 3-digit number), then delayed serial recall decreases precipitously over filled intervals of less than 20 s (Peterson & Peterson, 1959).

Bjork's discussion focussed on the data reported by Hellyer (1962), who presented participants with a consonant trigram, followed by 0, 1, 2, 4, or 8 repetitions, after which there was a filled retention interval of 3, 9, 18, or 27s. The findings from a near-replication of Hellyer (1962) appear in Figure 1 (Ward, Cortis Mack, Doherty, Knight, & Loaiza, 2019). In our experiment, participants were presented with three words, and were then prompted by the computer to rehearse the words aloud by re-presenting the stimuli at regular intervals. Participants were then either tested immediately (such that the retention intervals was 0s) or they were tested after an interval filled with a backwards-counting task (such that the retention intervals were 3, 6, 9 or 18s). The upper panel shows serial recall using the Brown-Peterson scoring system in which correct recall is credited only if the complete three-item sequence is recalled in the correct order. The middle and lower panels show the recall performance when participants can gain credit for recalling every list item that is remembered in the correct order (serial recall, Figure 1B) and every list item that is remembered regardless of the recall order (free recall, Figure 1C).

--Figure 1 about here—

Like Hellyer (1962), our data in Figure 1A show that (a) the delayed serial recall of a three-item sequence decreased with increasing retention interval and (b) recall benefitted from increased repetitions prior to the filled interval. Moreover, a comparison between the three panels show that although recall improved somewhat when credit was additionally given to partially correct lists, there were nevertheless very similar effects of repetitions and retention interval using all three measures, indicating that the manipulations of repetitions and retention interval affected access to item as well as order information.

Our discussion of the Brown-Peterson task has so far assumed that the backwards counting that occurs during the filled retention interval effectively eliminates rehearsal, but two further points need to be made regarding this assumption. First, Kroll and Kellicut (1972) have shown that rehearsal may still take place during a nominally filled retention interval. In their first study, they showed that successful recall of a letter trigram decreased at the longer delays when participants were required to continually subtract 7 rather than to subtract 3 from a 3-digit number during the filled retention interval. In their second experiment, they examined whether the consonant trigram was covertly rehearsed by asking participants to press a button if they thought of the letters in any way during the 9s retention interval. They again found that recall accuracy of the trigram was superior in the subtract 3 rather than the subtract 7 conditions, and there were more button presses in the former relative to the latter. When the authors plotted recall as a function of button presses they found a positive linear relationship between rehearsal and recall for both counting tasks.

Second, Vallar and Baddeley (1982) also conducted a Brown-Peterson task and varied the filler task that was performed during the retention interval. Specifically, participants were tested after 0, 5, or 15s and during the retention intervals, participants (i) performed repeated manual tapping, (ii) continuously subtracted 3 from a 3-digit number, or (iii) performed articulatory suppression (repeating “the” continuously without interruption). Performance following an immediate test (0s) was equivalent and above 95% correct in all three conditions. Consistent with prior Brown-Peterson findings, there was no forgetting for the 5s and 15s intervals when they were filled with the manual tapping (a condition where articulatory rehearsal could occur), and recall accuracy decreased substantially with the subtraction-filled delay (to 45% following 5s and to 30% following 15s filled delays, respectively). Critically, there was only the most modest forgetting in the articulatory suppression condition (accuracies were at 95% and 90% following 5s and 15s filled delays, respectively). This shows that the backward counting task doesn’t only prevent rehearsal: the articulatory suppression prevented rehearsal but this alone was insufficient to cause dramatic short-term forgetting. One interpretation is that there may be other non-articulatory methods for maintaining items in working memory (e.g., attentional refreshing) that can be used during articulatory suppression but which can be less readily used during a more attentionally-demanding task such as backward counting.

Finally, observant readers will have noticed that our discussion of the Brown-Peterson task has side-stepped any discussion of the nature of the short-term forgetting that is observed in a filled retention interval. Brown (1958) attributed the forgetting to trace decay – forgetting owing to the passage of time. An important proposed function of rehearsal was therefore to offset the effects of decay. For many, evidence that short-term forgetting is minimal on the first trial provides clear evidence against decay (Keppel & Underwood, 1962). However, Ricker, Vergauwe & Cowan (2016) are unconvinced by such findings. They argue that in the only experiment performed by Keppel and Underwood (Experiment 1) that did not suffer from ceiling effects, the rate of forgetting on trial 1 is no different to those on trial 2 and trial 3, clearly in agreement with Brown (1958)'s theory of trace decay.

As we will see, the concept of trace decay remains in many leading accounts of short-term (Atkinson & Shiffrin, 1968, 1971) and working memory (e.g., Baddeley, 1986; Barrouillet, Bernardin, & Camos, 2004; Cowan, 1988, 1995, 1999) and it is still invoked to explain short-term forgetting (e.g., Barrouillet, Portrat, Vergauwe, Diependaele, & Camos, 2011). It is worth pointing out that proponents of trace decay have sometimes needed a new post-stimulus memory mechanism, such as short-term consolidation, to help explain their findings (Ricker & Cowan, 2014).

Overall, from the Brown-Peterson literature, there is considerable evidence in support of a positive role of rehearsal on later recall. Not only is there little or no forgetting in the Brown-Peterson task during an unfilled interval (which could be attributed to a maintenance role of rehearsal), but rehearsing the stimuli prior to a filled retention interval greatly enhances later recall above non-rehearsed levels (e.g., Hellyer, 1962; Ward et al., 2019).

Rehearsal in the Immediate Free Recall Task

In the immediate free recall (IFR) task, participants are typically presented with reasonably long lists of unrelated items, one at a time, and at the end of the list, they must try to recall as many of these items as they can in any order that they wish. For example, Murdock (1962) presented participants with lists of between 10 and 40 words at presentation rates of 1s or 2s per item. He found recall advantages for the first few list items and the last few list items (recall advantages known as the primacy effect and the recency effect, respectively), resulting in a J-shaped or U-shaped serial position curve (see also Deese & Kaufman, 1957; Jahnke, 1965). Participants tended to recall a higher proportion of the list with slower presentation rates and at shorter list lengths. Some of these features of IFR are illustrated in Figure 2, which shows data from an IFR experiment in which participants were presented with lists of 10, 20, or 30 words at a rate of 3s per item, using the overt rehearsal method (Ward, 2002). Figure 2A clearly show the characteristic primacy and recency effects in IFR as well as clear list length effects.

--Figure 2 about here—

Rehearsal was proposed to serve two functions in classic dual-store theories of IFR (e.g., Atkinson & Shiffrin, 1968, 1971; Lehman & Malmberg, 2013; Raaijmakers & Shiffrin, 1981; Waugh & Norman, 1965), which proposed separate short-term memory (STS) and long-term memory stores (LTS). The STS was assumed to be of very limited capacity, retrieval from STS was assumed to be fast and efficient, and the nature of storage in STS was temporary in nature. Rehearsal was argued (1) to maintain items in the STS and (2) to increase the associative strength in a more durable LTS. Retrieval from LTS involved a probabilistic search and this was affected by the relative and absolute strength of the items (relative to competing items in the list).

In the original Atkinson and Shiffrin model (1968), it was assumed that items were lost from STS due to displacement and trace decay. This model could explain the Brown-Peterson forgetting task: the precipitous decline with a filled retention interval could be explained by trace decay[1], the Hellyer (1962) data could be explained through the repeated rehearsals strengthening traces in LTS (leading to raised asymptotes), and the Keppel and Underwood data could be explained by near-perfect recall from LTS on trial 1 (prior to the effects of proactive interference from previous recall trials, see Healy and McNamara, 1996).

The Atkinson and Shiffrin model could also explain the data from Murdock (1962). The recency effects were argued to reflect the direct output from STS, and the primacy effects were argued to reflect the possibility that the early list items received greater numbers of rehearsals than later list items. The model could also explain the effects of presentation rate, if one assumed that items presented at slower rates would receive more rehearsals and so be encoded more strongly in LTS, whereas the list length effects could be explained if one assumed that the relative strength of a given item would be greater in a shorter list (where there would be fewer competitors).

Subsequent work appeared to be consistent with a rehearsal-based account of the primacy effect in IFR. Rundus (1971) presented participants with lists of 24 unrelated words

[1] It is interesting to note that the trace decay that allowed the Atkinson and Shiffrin model (1968) to explain the forgetting in the Hellyer data was not assumed by Waugh and Norman (1965). Rather, they assumed that items from STS were lost through displacement not trace decay. Waugh and Norman had presented participants with long strings of digits and at unpredictable points, had asked participants to recall the item that followed a re-presented probed digit. Accuracy declined steeply with numbers of intervening items (a finding that they likened to the recency effects in IFR, above). However, they found that there was only a small change in performance (a slight improvement after 5 intervening items) on the digit probe task when the rate was increased from 1 item per second to 4 items per second, and they reasoned that there should have been a more substantial increase in recall given a four-fold increase in presentation rate if items were assumed to decay with time.

at a slow rate for IFR and asked participants to say aloud the words that came to mind during the presentation of the lists (the procedure known as the overt rehearsal method, Rundus & Atkinson, 1970). Although Rundus' (1971) primary focus was on the number of rehearsals that items receive, it is worth reiterating that Rundus was explicit in stating that participants' rehearsals need not only arise through rote repetition but may also be generated following a variety of strategies including the forming of inter-item associations, the use of mnemonics, visual imagery, and organisation.

Rundus found large primacy and recency effects, and consistent with a rehearsal-based account of the primacy effect, he found that the earlier list items received far more rehearsals than later list items. This finding is shown in the Ward (2002) data in Figure 2B. By contrast, the recency items were well recalled but received the fewest rehearsals. These patterns of rehearsal may help explain the data from a test of final free recall (Craik, 1970) in which participants are tested at the very end of the experiment to try to recall any words from any lists that they had encountered during the experiment. Since final free recall relies solely on retrieval from LTS, recall from all serial positions will be predicted by the frequency of rehearsals. Consistent with these predictions, the final free recall curves show extended primacy effects but negative recency effects.

A rehearsal-based explanation of the primacy effect in IFR was further strengthened by manipulations designed to modify the opportunities to rehearse. Fischler, Rundus and Atkinson (1970) and Glanzer and Meinzer (1967) presented participants with lists of words for IFR but asked the participants to use a fixed-rehearsal strategy in which they rehearsed each word only whilst it was being presented. They found that by using this strategy, the number of rehearsals was kept constant for all list items and the primacy effect was greatly attenuated.

Developments in dual-store accounts of free recall further specified the role of rehearsal. The Search of Associative Memory (SAM) model by Raaijmakers and Shiffrin (1981) detailed how rehearsal served multiple functions: it increased the associative strength in LTS between an item and the list context, it increased the associative strength in LTS between an item and other co-rehearsed items, and it increased the associative strength in LTS between an item itself.

Not all studies examining rehearsal in IFR have supported the positive relationship between rehearsal and final recall. A number of studies contested the claim that later recall was improved by increasing the number of rehearsals (e.g., Craik & Watkins, 1973; Glenberg, Smith & Green, 1977; Jacoby, 1973; Rundus, 1977; Woodward, Bjork, & Jongeward, 1973; for a more detailed consideration, see later section). In these studies, the experimental procedure manipulated the number of rehearsals that were afforded to different words presented during an experimental session. These extra rehearsals tended to be massed together as a block (or within a list). In a final free recall task, participants did not show systematic recall advantages for words that were rehearsed more frequently.

Other studies have shown that factors in addition to the frequency of rehearsal affect later recall. It had previously been shown that the mnemonic benefits of repeating a presented item in the list increased when the spacing or lag between the repetitions increased (e.g., Madigan, 1969; Melton, 1970). Subsequent studies showed that improved recall also arose from increased numbers of rehearsals that were distributed throughout the presentation period (e.g., Tan & Ward, 2000; Modigliani & Hedges, 1987) and that were last rehearsed to more recent list positions (e.g., Brodie, 1975; Brodie and Murdock, 1977; Brodie & Prytulak, 1975; Tan & Ward, 2000; Ward, 2002; Ward & Tan, 2004). These studies showed that as well as plotting the serial position curves as a function of when the item was presented (termed the nominal order), it can also be beneficial to plot the serial position curves as a function of the rank order in which the items were last rehearsed (termed the functional order), or simply as a function of when a particular word was last rehearsed.

An illustration showing the importance of recency of rehearsal can be seen in Figure 2. Figures 2C and 2D re-plot the same data from Figure 2A by each item's functional order (Figure 2C) and by when each item was last rehearsed (Figure 2D). As one can see, re-plotting suggests that the recency effect is not limited solely to the most recent items, but rather the effects of recency are present throughout the entire list (with a possible exception of the first item). In these plots, the primacy effect is greatly attenuated, suggesting that the first few items benefit from being rehearsed toward the end of the list. Early list items that are not later rehearsed are relatively poorly recalled. Tan and Ward (2000) argued that episodic memory could be represented as a continuum, that list items (and their rehearsals) were accessed according to a recency-based function, and that the number, recency, and distribution of rehearsals were all important in predicting recall.

One criticism of rehearsal-based explanations is that they tend to show correlations between rehearsed words and recalled words, but the relationship need not be causal (e.g., Souza & Oberauer, 2018). If some words are more familiar, more distinctive, or otherwise more memorable, then these memorable words may be both rehearsed more often and recalled more often. The positive correlation between rehearsal and recall would then stem from the idiosyncratic nature of the words themselves rather than a causal link between rehearsals causing improved recall. In response, a number of studies have presented schedules of "rehearsals" to new participants. These "rehearsals" were either generated by computer algorithms (e.g., Murdock & Metcalfe, 1978; Tan & Ward, 2000) or had been previously recorded by one group of participants and then the schedules applied to a second group of participants, sometimes studying different words (e.g., Brown, Della Sala, Foster & Vousden, 2007; Tan & Ward, 2000). The subsequent patterns of recall showed similar benefits for these yoked rehearsals, suggesting that it is the pattern of repetitions rather than solely the identity of the originally rehearsed items that conveys the recall benefit.

Laming (2006, 2008, 2009, 2010) has considered further the relationship between overt rehearsals and recalls. Central to Laming's hypothesis is the proposal that the "process that generates recalls is the same process that generates rehearsal, subject only to the restriction that recalls are seldom repeated." (Laming, 2006, p.1146)[2]. According to Laming, the accessibility of earlier rehearsals (for later rehearsals and then recall) decreases with increasing number of intervening items, but once an item is recalled, the next item in the sequence can be retrieved with an enhanced probability (see also, the temporal context model of Howard & Kahana, 2002). Both Laming (2006) and Ward, Woodward, Stevens & Stinson (2003) have shown that the output order in IFR was greatly affected by the co-rehearsals during presentation: participants were highly likely to output successive words that were last rehearsed in the same order (Laming, 2006) or that were last co-rehearsed in the same interval between the words (Ward et al., 2003). Indeed, over a number of related papers, Laming has shown that the observed schedules of recall in IFR can be predicted (at least to some extent) by the full schedules of rehearsals.

If recall is related to rehearsal as Laming (2006) claimed, then one might question why participants in IFR often rehearse in a cumulative, forward-ordered manner (at least early in the presentation of the list), but recall in a highly recency-based manner (e.g., Tan & Ward, 2000). Relevant data pertaining to this issue were provided by Ward, Tan and Grenfell-Essam (2010) who showed that participants tend to initiate IFR with one of the last list items with longer lists of 10 or more words (as already established by Hogan, 1975; Howard & Kahana, 1999; Laming, 1999), but when the list length is shorter than 5 words, the participants tend to initiate IFR with the first list item. That is, participants tend to recall a

[2] This idea can be traced back to Metcalfe (1975, cited in Laming, 2006) and Laming's analyses rely heavily upon the data from Murdock and Metcalfe (1978).

short list presented for IFR, such as “book, house, fish”, in an “immediate serial recall (ISR)-like” manner (that is, they recall: “book, house, fish”), even though serial recall was not a task requirement. Grenfell-Essam, Ward and Tan (2013) examined the effects of articulatory suppression and presentation rate on this tendency to initiate IFR of short lists of words with the first item and showed that this tendency was unaffected by fast rates of 2 words per second, and this tendency was still present (albeit somewhat reduced) under articulatory suppression. The findings of Grenfell-Essam et al. (2013) suggest that at short lists, even in the absence of rehearsal, participants show a preferred tendency to output in forward serial recall; and this forward ordering of recall is like the forward ordering of rehearsal that occurs early in the presentation of a list (cf. Laming, 2006).

Overall, from the IFR literature, there is considerable evidence in support of a positive role of rehearsal on later recall. Within dual-store accounts of IFR, rehearsal is assumed to be a controlled process within STS, and rehearsal results in the strengthening of associations within LTS resulting in greater recall. Within unitary memory accounts, episodic memory is considered to be a continuum from the most recent past to the distant past. Rehearsal is considered to be a mini-recall using the same retrieval mechanism(s) as those that are used at test. A theme of this work is that rehearsal imparts a trace into memory much like a repetition of a study event, with recall dependent upon the number, the distribution, and the recency of the rehearsals on the list.

Rehearsal in the Immediate Serial Recall (ISR) Task

In the ISR task, participants are typically presented with lists of 5-9 items and they must try to recall the items in *exactly the same order* as they had been presented. Performance on the ISR task is nearly perfect for very short lists of 3 or 4 items, but as the list length increases, so performance quickly deteriorates, resulting in a bowed serial position curve with highly extended primacy effects and reduced recency effects.

Early studies proposed that ISR was particularly sensitive to speech-based variables (e.g., Baddeley, 1966; Murray, 1967) and it was assumed that subvocal rehearsal was an important way of maintaining verbal information in a temporary and highly activated state (Baddeley & Hitch, 1974). One of the most influential and widely-cited accounts of ISR is the Phonological Loop model of working memory (e.g., Baddeley, 1986, 2012; Baddeley & Hitch, 1974, 2019). The phonological loop was proposed to consist of a phonological short-term store coupled with an articulatory loop. The phonological nature of the store was responsible for the phonological similarity effect, the recall advantage in ISR for dissimilar-sounding items relative to similar-sounding items (e.g., Baddeley, 1966). Silently presented visual-verbal stimuli were first articulated and then recoded into the phonological store, whereas auditory items entered the phonological store directly. Items within the phonological store were assumed to decay with time, unless they were rehearsed in the articulatory loop (Baddeley, 1986).

Evidence that rehearsal offset trace decay was provided by the word length effect, the advantage for shorter words over longer words in ISR. In one study, Baddeley, Thomson, and Buchanan (1975) examined ISR for lists of 5 words that each contained 1, 2, 3, 4, or 5 syllables. They found that memory span decreased as the number of syllables in each word in the list increased; and the decline in span directly correlated with the increased time to read the words. The authors also observed a duration-based disyllabic word length effect: a recall advantage for 2-syllable words that took less time to pronounce relative to 2-syllable words that took a longer time to pronounce. In a number of subsequent studies, differences in the memory span in ISR for different materials were directly linked to the differences in the time taken to verbally rehearse (e.g., Ellis & Hennelly, 1980; Schweikert & Boruff, 1986;

Standing, Bond, Smith & Isely, 1980) and differences between different participants' memory spans were attributed to differences in rehearsal rates (e.g., Hulme, Thomson, Muir, & Lawrence, 1984; Naveh-Benjamin & Ayres, 1986). It was argued that lists of shorter words could be rehearsed more effectively than lists of longer words and so were less affected by trace decay. The importance of the word length effect was heightened by comments such as those of Cowan (1995, p.42) who claimed that the word length effect was "perhaps the best remaining solid evidence in favour of temporary memory storage".

When rehearsal was prevented through articulatory suppression, ISR performance decreased, and this was assumed to be because the contents of the phonological store were decaying, but owing to the concurrent articulation, they could not be reactivated through rehearsal. Moreover, word length effects and phonological similarity effects were abolished when written words were presented under articulatory suppression: the concurrent articulation was assumed to prevent the memoranda from even entering the phonological store (Baddeley, Lewis & Vallar, 1984).

Initial developmental data also appeared to be consistent with a rehearsal-based account of ISR. Gathercole and Baddeley (1989) conducted a longitudinal study to examine whether the phonological loop might be involved in language acquisition. They developed the nonword repetition task in which children heard a set of 40 spoken nonwords of varying syllable length and complexity and had to immediately repeat back the nonwords. Gathercole and Baddeley tested the vocabulary skills of 104 children within two months of entering primary school (between the ages of 4 and 5) and then retested these skills one year later. They also measured each child's nonverbal intelligence, nonword repetition and reading skills on both occasions. They found that children's nonword repetition was highly correlated with vocabulary at age 4 ($r = .53$) and age 5 ($r = .49$), and the nonword repetition score at age 4 predicted vocabulary at age 5 ($r = .57$). These correlations remained significant even after the effects of other psychometric values such as nonword intelligence, standard reading rate and conventional digit span were partialed out. They argued that the data were consistent with the phonological loop being involved in language acquisition. Further developmental data suggested that children's accuracy on ISR increase markedly between the ages of 5 and 9, a time at which they were argued to develop phonological memory skills. Children tend to show spontaneous evidence of rehearsal around the age of 7, and studies show that it is around this age that children first demonstrate significant phonological similarity effects and word length effects (e.g., Gathercole, 1998; Henry, Messer, Luger-Klein, & Crane, 2012).

However, there has been on-going controversy over whether short-term forgetting is caused by trace decay offset by rehearsal (Brown & Hulme, 1995; Lewandowsky, Oberauer, & Brown, 2009; Nairne, 2002; Neath & Brown, 2006; Neath & Nairne, 1995; Oberauer & Lewandowsky, 2008; Souza & Oberauer, 2018). Here we consider briefly three lines of evidence against a trace decay interpretation (but for a more detailed critique of decay, see Lewandowsky & Oberauer, 2015): (i) issues with the word length effect, (ii) issues concerned with lack of forgetting when rehearsal was suppressed, and (iii) reinterpretation of developmental data.

First, there have been concerns about the generalizability of word length effects because there have been difficulties replicating word length effects using different sets of words. Considering the duration-based disyllabic word length effect, Baddeley et al. (1975) had originally shown that ISR performance on lists of five short-duration disyllabic words (taken from the set: bishop, pectin, ember, wicket, wiggle, pewter, tipple, hackle, décor, and phallic) was superior to ISR performance on lists of five long-duration disyllabic words (taken from the set: Friday, coerce, humane, harpoon, nitrate, cyclone, morphine, tycoon, voodoo, and zygote). However, Caplan, Rochon, and Waters (1992) observed a reversal of the word length effect with different sets of disyllabic words. In their data, long-duration disyllabic words were better recalled than short-duration disyllabic words, in direct contradiction to the duration-based disyllabic word length effect. Moreover, Lovatt, Avons and Masterson (2000) showed a disyllabic word length effect when Baddeley et al.'s sets of words were used, but when different sets of words were used then long words were sometimes better than short words and sometimes there was no difference. The importance of the exact set of words has been most clearly demonstrated by an experiment by Neath, Bireta and Surprenant (2003), who examined disyllabic word length effects in ISR using the word sets from all three papers (Baddeley et al., 1975; Caplan et al., 1992; Lovatt et al., 2000). Neath et al. demonstrated that different word sets do indeed produce different results: bad news for a general principle of time-dependent word length effects. Considering the multisyllabic word length effect, Jalbert, Neath, Bireta, and Surprenant (2011) have argued that previous studies had confounded word length with orthographic neighborhood size. When orthographic neighborhood size was controlled between a set of 1-syllable and 3-syllable words, there was no word length effect. Guitard, Gabel, Saint-Aubin, Surprenant and Neath (2018) further argued that if one controlled a full range of factors (including concreteness, imageability, familiarity, word frequency and orthographic neighborhood size) then the word length effect between 2- and 3-syllable words disappears.

Second, a number of studies failed to find evidence of trace-decay following a filled delay (e.g., Lewandowsky, Duncan, & Brown, 2004; Lewandowsky, Nimmo, & Brown, 2008; McFarlane & Humphreys, 2012; Oberauer & Lewandowsky, 2008), and we have already reviewed that there was little short-term forgetting in the Brown-Peterson task when articulatory suppression was performed during a filled delay (e.g., Vallar & Baddeley, 1982).

Third, developmental data (Alloway, Gathercole, & Pickering, 2006; Gathercole, Pickering, Ambridge, & Wearing, 2004) showed that memory performance increased linearly with age (from 4.5 to 10.5 years) for a range of memory measures, including both verbal and spatial tasks. These data do not provide clear evidence of a phonological-specific improvement around the age of 7 (Lewandowsky & Oberauer, 2015).

The previous paragraphs reviewed evidence for the role of rehearsal in ISR using indirect measures (such as the word length effect) and considered the proposed functions of rehearsal in ISR: to maintain items in working memory by offsetting the effects of trace decay and to help in language acquisition. However, the ISR data we have considered so far have been collected using only indirect manipulations of rehearsal: factors other than the speed of rehearsal could be affecting the ISR of words of different length, articulatory suppression may do more than simply prevent rehearsal, and the relationship between ISR and language acquisition was only correlational in nature.

More direct evidence for a role for rehearsal in ISR comes from studies using the overt rehearsal methodology (e.g., Bhatarah, Ward, Smith & Hayes, 2009; Souza & Oberauer, 2018; Tan & Ward, 2008). Tan and Ward (2008) presented two groups of participants with lists of 6 words for ISR. One group studied the list items under visual silent conditions, whereas the rehearsal group studied the list items using the overt rehearsal methodology. Both groups saw lists presented at fast (1s per word), medium (2.5s per word) and slow (5s per word) rates. Tan and Ward (2008) found that ISR increased with slower presentation rates in both groups, and in both groups, the data showed characteristic extended primacy effects in ISR at all presentation rates. The ISR data for the overt rehearsal group are shown in Figure 3A.

--Figure 3 about here—

An examination of the overt rehearsals revealed that participants tended to rehearse only the currently presented item at fast presentation rates (fixed rehearsal). At slower presentation rates, such as the slow 5s per item presentation rate shown in Figure 3B, participants tended to perform cumulative rehearsal (black columns) for the first four serial positions. Figure 3C shows that there were positive correlations between accuracy in ISR and the maximum length of sequence that was rehearsed during the trial. Those participants who at slower rates rehearsed longer sequences were more accurate at ISR, whereas those participants who at slower rates rehearsed shorter sequences were less accurate at ISR. The Tan and Ward (2008) data suggest that cumulative rehearsal is not necessary for ISR at fast rates, but at slower rates, rehearsal enhances later recall by more than simply maintaining the list items (cf. Hellyer, 1962).

Bhatarah, Ward, Smith and Hayes (2009) replicated the Tan and Ward (2008) findings and additionally showed that the patterns of rehearsals observed in ISR were very similar to those observed in IFR. Indeed, a growing body of research suggests that participants perform IFR of short lists in an ISR-like manner (see Ward et al., 2010). For example, Bhatarah et al. (2009) showed that ISR and IFR were somewhat similarly affected by manipulations of word length and articulatory suppression; Bhatarah, Ward and Tan (2008) and Grenfell-Essam and Ward (2012) showed that words in ISR and IFR were encoded in a similar manner such that performance on the two tasks was relatively unaffected by manipulations of test expectancy; and Spurgeon, Ward and Matthews (2014) showed that ISR and IFR were similarly affected by manipulations of phonological similarity and articulatory suppression. Thus, in both IFR (Tan & Ward, 2000) and ISR (Tan & Ward, 2008) there appear to be positive mnemonic effects of rehearsal on subsequent recall.

One interesting feature of the overt rehearsal schedules generated by Tan and Ward's (2008) participants is that they rarely performed cumulative rehearsal of all six items during ISR, but more often rehearsed only around 4 items in forward order. One might wonder why participants do not extend this rehearsal strategy so that they rehearse the entire list? Jarrold

(2017) has argued that serial rehearsal may be most effective when the number of items to be rehearsed is subspan. Using an indirect method, Jarrold, Tam, Baddeley and Harvey (2010) presented participants with sequences of six words for ISR. Unlike conventional ISR studies, in Jarrold et al.'s lists, participants always had to engage in one block of 18s of distractor activity (involving making verbal or visuo-spatial decisions to presented letter pairs) that could occur before or after any one of the six items. When participants performed visuo-spatial decisions during the processing interval (and so were able to rehearse) the accuracy of ISR for the first word was highest when the processing task was after one of the first 2-4 items, but when the processing task was after the fifth or sixth word, the accuracy of ISR for the first word declined. One interpretation is that participants were able to rehearse sequences of up to 4 items throughout the 18s processing task, and that effective rehearsal broke down when participants had to rehearse 5 or 6 items (Jarrold, 2017).

Two very recent studies also demonstrate that ISR performance does not improve when participants are required to rehearse more than 4 items. First, Souza and Oberauer (2018) have questioned whether there is a causal link between increased articulatory rehearsal and increased serial recall. They replicated the Tan and Ward (2008) study, performing ISR with overt rehearsal. Like Tan and Ward (2008), they showed that ISR improved at slower presentation rates and they also found a positive correlation between ISR accuracy and the maximum length of sequence that was rehearsed in cumulative forward order. However, in a critical manipulation, they also compared rehearsal and recall between one group of participants who were free to rehearse aloud however they so wished (standard overt rehearsal instructions) and a second group of participants who were instructed to perform cumulative forward rehearsal at their maximum rate (to the best of their abilities). Those participants who were explicitly instructed to perform cumulative rehearsal at their maximum rate did rehearse the words more frequently than those performing standard overt rehearsal, and they also rehearsed significantly longer sequences of items. However, the increased rehearsal did not lead to any increase in ISR performance. Moreover, in a third experiment, Souza and Oberauer (2018) presented three groups of participants with sequences of 6 words for ISR. One group was presented with the words at a fast rate of 1s per word, a second group was presented with the words at a slow rate of 5 s per word, but performed articulatory suppression (continuously saying “babibu”), and a third group was also presented with the words at a slow rate of 5 s per word, but had to read aloud the rehearsals of a yoked participant. Souza and Oberauer found no serial recall advantage for the yoked rehearsal group over the articulatory suppression group (although there was a recall advantage for the yoked rehearsal group over the articulatory suppression group using free recall scoring, as reported in Souza & Oberauer, 2020). These findings have led Souza and Oberauer (2018) to argue against a causal link between articulatory rehearsal and ISR (see also Lewandowsky & Oberauer, 2015; Oberauer, 2019; Souza & Oberauer, 2020).

Second, a recent series of studies by Barrouillet, Gorin and Camos (in press) has provided a possible explanation for Souza and Oberauer’s (2018) failure to find a recall advantage in ISR when participants were instructed to perform cumulative rehearsal to the maximum of their ability. In each of three experiments, Barrouillet et al. (in press) presented two groups of participants with a series of between 3 and 5 blue letters followed by a series of between 1 and 6 black letters for ISR. One group of participants always received instructions that the authors called *the maxispan procedure*: these participants were invited to perform cumulative rehearsal of the blue letters throughout the presentation of all the black letters until they received the cue to perform ISR of all the letters (recalling the blue letters first, followed by the black letters). Performance on the maxispan instructions was always contrasted with a second simple span group who received standard ISR instructions: this group received similar sequences of blue then black letters, but they were told to perform ISR and ignore the colour of the letters. Barrouillet et al. found that the recall of the blue letters was almost perfect in the maxispan group and far superior to the recall of the blue letters in the simple span group. In addition, Barrouillet et al. found that participants in the maxispan group also recalled more black letters than did participants in the simple span group. Interestingly, Barrouillet et al. found that increasing the number of blue (rehearsed) letters from 3 to 4 had little effect on the recall of the later black letters, but increasing the blue (rehearsed) letters to 5 did impact later black letters. Far from arguing against a causal role of rehearsal in ISR, Barrouillet et al. argued for a critical role for the rehearsal of a limited subset of items in ISR. They argued that participants often try to rehearse too many stimuli when performing ISR (Souza & Oberauer, 2018) and so overload their phonological loop resulting in poorer ISR performance.

In summary, rehearsal has been proposed to play a major role in ISR to maintain items in working memory to offset the effects of trace decay, but there are concerns over whether rehearsal is the only factor being manipulated using indirect manipulations, and whether forgetting is primarily driven by trace decay. Moreover, using indirect manipulations one cannot be certain of what was actually rehearsed on any given trial. By contrast, using the overt rehearsal method, one can be more certain of the schedule of rehearsals associated with each retrieval attempt. Using overt rehearsal, there is consistent evidence that ISR performance increases at slower rates, such that when participants are free to rehearse however they so wish, their ISR accuracy at slower rates exceeds their performance at faster rates and their ISR accuracy correlates with the degree to which they performed cumulative forward-ordered rehearsal. It may be that the serial recall advantages associated with rehearsal are limited to rehearsing only a subset of words (Barrouillet et al., in press; Jarrold, 2017) because increasing cumulative rehearsal beyond sequence lengths that participants would themselves choose leads to no further increase in ISR (Souza & Oberauer, 2018).

Rehearsal in the complex span task

In recent years, complex span tasks have become some of the most frequently used methods in working memory research. This may be because performance on complex span tasks correlates more strongly (than performance on ISR and IFR) with measures of higher order cognition such as intelligence (e.g., Conway et al., 2005; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004; Oberauer et al., 2005). In complex span tasks, participants must maintain a sequence of to-be-remembered items (e.g., letters, words, digits or visuo-spatial locations) whilst also engaging in intervening processing tasks (e.g., mental arithmetic, shadowing, or choice reaction time tasks). A rehearsal strategy is often reported in these tasks to maintain items (e.g., Dunlosky & Kane, 2007; Turley-Ames & Whitfield 2003), and when participants are trained to rehearse the memoranda in cumulative forward order they tend to perform more accurately on the complex span task (but this could simply reflect the additional time spent at encoding, Lewandowsky & Oberauer, 2015).

A leading account of complex span, the Time-Based Resource-Sharing (TBRS) model (Barrouillet & Camos, 2012, 2015; Barrouillet, Bernardin & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007) assumes that memoranda will suffer from trace decay if they are not maintained in working memory. Two maintenance mechanisms have been proposed (for a review, see Camos & Barrouillet, 2014): articulatory rehearsal (a domain-specific mechanism for verbal stimuli which is similar to the idea of the Phonological Loop reviewed in models of ISR), and *attentional refreshing* used to maintain verbal and non-verbal stimuli by focusing attention on the decaying traces of the memoranda (see also Camos, 2015, 2017; Raye, Johnson, Mitchell, Greene, & Johnson, 2007).

Ward et al. (2019) have presented direct evidence that participants do rehearse in complex span tasks when the opportunity presents itself. They examined participants' ability to perform complex span using the overt rehearsal method. Participants were presented with lists of between 1 and 6 words that they were asked to recall in serial order. Following each word in the list was a series of visuo-spatial choice reaction time tasks (using button press responses) that are widely used to distract attention without preventing overt rehearsal. The serial position curves for the six different list lengths are presented in Figure 4A. At shorter list lengths, the serial position curves were relatively flat, but when there were 5 or 6 words in the list, extended primacy effects were observed.

--Figure 4 about here—

As can be seen from Figure 4B, the patterns of rehearsal were similar to those observed by Tan and Ward (2008) using ISR: there was clear evidence of cumulative forward-ordered rehearsal (at least for the first 3-4 words) and as can be seen from Figure 4C, there was a positive correlation between accuracy in serial recall and the length of the maximum sequence of words that was rehearsed.

Concerning next the second maintenance process, the metric of *Cognitive Load* (CL) has been used to calculate the proportion of time within an intervening processing task that participants must engage in attentionally-demanding processes. Formally, the metric of *Cognitive Load* (CL) is calculated by the equation, CL = a N / T, where *T* is the total time during which the processing task takes place, *N* is the number of attentionally-demanding retrievals that each capturing attention for time, *a*. Consistent with the TBRS model, a benchmark finding (Oberauer et al., 2018) is that there is a negative linear relationship between working memory span and Cognitive Load (e.g., Barrouillet et al., 2004, 2007; Langerock, Vergauwe, & Barrouillet, 2014; Vergauwe, Langerock, & Barrouillet, 2014).

Evidence in support of two distinct mechanisms comes from Camos, Lagner and Barrouillet (2009) who showed that there were additive effects of interrupting rehearsal (by requiring oral rather than button press responses to the processing task) and performing an attentional demanding task (by making an odd/even choice reaction time task rather than a digit detection simple reaction time task). In addition, Camos, Mora and Oberauer (2011) have argued that participants can use both articulatory rehearsal and attentional refreshing in complex span tasks, and they use these strategies adaptively. They presented phonologically similar and phonologically dissimilar lists for complex span and varied the ability to engage in attentional refreshing (by using either a simple reaction time or choice reaction time task). Under higher attentional loads, participants were more reliant on articulatory rehearsal and showed stronger phonological similarity effects. These were present using both choice and simple reaction time processing tasks when participants were told to engage in rehearsal and were absent using both choice and simple reaction time processing tasks when participants were told to engage in refreshing.

Barrouillet et al. (in press) have recently argued that since the two mechanisms are each assumed to have a capacity of about 4 items, that they should operate together optimally in the maxispan procedure when the articulatory rehearsal encodes up to 4 blue list items and the attentional refreshing mechanism encodes up to 4 later black list items. As has already been reviewed, participants' ISR spans increase using the maxispan procedure relative to standard ISR.

Recently, Souza and Oberauer (2020) have further examined the role of rehearsal in simple and complex span tasks. In Experiment 1, four groups of participants were presented with lists of six words for serial recall at both medium and slow presentation rates. Two groups performed ISR (or simple span) without any distractors (one group rehearsed aloud, the other silent) and two groups performed complex span tasks with visuo-spatial distractors (one group rehearsed aloud, the other silent). In the distractor task, participants saw a horizontal line that was presented above or below a pair of separated squares and had to decide whether or not the line fitted within the squares and respond by pressing one of two keys. Consistent with Tan and Ward (2008) and Ward et al. (2019), there were extended primacy effects in the serial position curves of both simple and complex span, and there were recall advantages in all four groups for words at slow relative to medium presentation rates. In the overt rehearsal condition, there was evidence of increased cumulative rehearsal at slow rates (particularly for simple spans) and there were positive correlations between mean serial recall accuracy and the average maximum sequence length that was rehearsed on a trial in all conditions (these correlations were higher in complex span relative to simple span).

The causal relationship between rehearsal and serial recall in simple and complex span was tested in Souza and Oberauer (2020, Experiment 2). Two groups of participants performed complex span task on sequences of 6 words at slow rates with visuo-spatial distractors. All participants performed an initial pre-test block of trials on complex span, a middle training block on ISR, followed by a test block of trials on complex span. The two groups differed in the instructions given ahead of the training block. The Cumulative Rehearsal group were told to rehearse in both the training block and test blocks in a cumulative forward order and given examples of cumulative rehearsal strategy. The control group received no such instructions. In the training and test blocks, participants in the Cumulative Rehearsal group rehearsed far more than in the control group, but despite far greater rehearsal there was no improvement in serial recall. Based on these and earlier findings, Souza and Oberauer (2018, 2020) argued against a causal relationship between rehearsal and serial recall in simple and complex span tasks.

In summary, rehearsal has been implicated as one of two maintenance mechanisms in TBRS, the leading account of complex span performance. Although there is much indirect evidence for rehearsal (based largely on the presence of phonological loop effects in complex span and articulatory suppression), there have been far fewer direct studies of rehearsal using the complex span tasks. From both the Ward et al. (2019) and the Souza and Oberauer (2020) overt rehearsal data sets, it is clear that participants do engage in articulatory rehearsal in complex span (when the processing task so allows), and the patterns of rehearsals in complex spans share similarities with those from ISR (or simple span). Complex span performance and cumulative rehearsal both increase at slower rates (Souza & Oberauer, 2020), and there are large positive correlations between complex span and the maximum sequence length that was rehearsed (Souza & Oberauer, 2020; Ward et al., 2019). Nevertheless, although participants appear to benefit from rehearsal when they are free to do so, they fail to benefit from instructions that encourage participants to rehearse more that what they would naturally choose to do (Souza & Oberauer, 2020).

Currently, there is no empirical evidence that examines this new finding, to see if complex span (like ISR) only benefits from the cumulative forward rehearsal of a limited subset of items. For example, it would be interesting to insert a longer period of distractor activity after one of the six words to see if, like in Jarrold et al.'s (2010) study of ISR, effective rehearsal breaks down when the length of the rehearsal sequence nears span. Similarly, it would be interesting to see whether the complex span task benefits from maxispan instructions (Barrouillet et al., in press). Nevertheless, the recent findings of Souza and Oberauer (2018, 2020) represent important data that must first be fully explained if a causal relationship between cumulative rehearsal and serial recall is universally accepted.

Summary of the role of rehearsal in in Short-term Memory and Working Memory Tasks

We have seen that rehearsal has been proposed to play a significant role in the Brown-Peterson task, the IFR task, the ISR (or simple span) task and the complex span task.

The positive role of rehearsal is perhaps least controversial in IFR: most researchers accept that rehearsal improves later recall. For some dual-store theorists, rehearsal is assumed to offset the effects of trace decay in STS and increase the strength of association within LTS. For other researchers, rehearsal is assumed to create multiple copies of the stimulus trace with recall affected by the frequency, recency and distribution of these rehearsals. It is noteworthy that Oberauer (2019) in a somewhat downbeat assessment of rehearsal mechanisms in working memory excluded the role of rehearsal in IFR from his review.

In other tasks in which responses are required in the correct serial order, such as the Brown-Peterson task, the ISR (or simple span) task and the complex span task, the evidence for a causal relationship between rehearsal and recall is far more contentious (e.g., Lewandowsky & Oberauer, 2015; Oberauer, 2019; Souza & Oberauer, 2018, 2020). One interpretation of these data is that rehearsal improves recall performance when the length of the rehearsed sequence is subspan (e.g., Hellyer, 1962; Jarrold, 2017). This limit may naturally occur in self-generated schedules of rehearsals (Tan & Ward, 2008), and can be encouraged through maxispan instructions (Barrouillet et al., in press) in which only a small subset of the first few items are rehearsed throughout the presentation of later items to the very end of the list. However, when participants must maintain longer sequences for an extended duration, or when participants are encouraged to rehearse to the maximum of their abilities, there is a danger that the act of rehearsal may introduce order errors, leading to minimal benefits with serial recall scoring.

The Role of Rehearsal and Long-term Episodic Memory

In the previous sections, we considered the effects of rehearsal at encoding and retention on tests that were either conducted immediately or after only a brief delay. These sections showed evidence for a mainly positive effect of rehearsal on subsequent recall, particularly in tests of IFR.

In the sections that follow, we turn our attention to the role of rehearsal in long-term episodic memory. We start by considering the relationship between rehearsal and long-term retention, and we then consider the relationship between rehearsals and repetitions in episodic memory tasks.

Maintenance and Elaborative Rehearsal and long-term learning

Early interest in the role of rehearsal in long-term retention was stimulated by the predictions of dual-store accounts of memory (e.g., Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). As reviewed earlier, these models proposed that rehearsal increased the strength of associations of the items in LTS leading to enhanced long-term retention. These predictions contrasted with those of Craik and Lockhart (1972) and Craik and Watkins (1973) who argued that there were two types of rehearsal (maintenance rehearsal and elaborative rehearsal). Processing that encouraged a deeper analysis of the stimulus or that enriched and elaborated the memory trace, or elaborative rehearsal, leads to an improvement in long-term retention, whereas merely passive repetition or shallow analysis of a trace, or maintenance rehearsal, does not lead to a long-term mnemonic benefit.

In Greene's (1987) review of the effects of maintenance rehearsal on memory, he summarized a number of empirical studies that appeared to show that there was not a clear positive relationship between frequency of maintenance rehearsal and subsequent long-term retention when measured by recall (e.g., Craik & Watkins, 1973; Glenberg, Smith & Green, 1977; Rundus, 1977; Woodward, Bjork & Jongeward, 1973), but there appeared to be greater effects in tests of recognition. However, the additional maintenance rehearsals in the studies reviewd by Greene were often massed rather than distributed throughout the list (cf. Tan &

Ward, 2000), such that the manipulations of the frequency of rehearsal did not benefit from being rehearsed to more recent positons, nor did they benefit from being rehearsed and encoded in such a variety of different study contexts.

Later interest in this issue has arisen following studies examining the mnemonic consequences of performing simple span (ISR) and complex span tasks on a later test of retention such as final free recall. For example, McCabe (2008) presented participants with short lists of 2-4 words for either a simple span (ISR) or a complex span (operation span) task. After three trials of each type of list length and task, there was a final free recall test of all 54 words. In the immediate tests, performance had been essentially perfect for ISR but was reduced in the complex span task. By contrast, in the final free recall test, performance was superior for those words that had been presented in the complex span task relative to those that had been presented for ISR.

McCabe argued that the improvement in delayed recall for those words that had been presented for complex span arose because participants used covert retrieval from LTS to maintain items during the complex span task (this covert retrieval had not been necessary during ISR). The additional covert retrieval in complex span provided distributed practice, and the beneficial effects of distributed retrieval practice were observed in the test of final free recall. This interpretation was strengthened by McCabe's (2008) observation that the final free recall advantage showed primacy effects within the complex span task - the early items were argued to be covertly retrieved more often than later items.

That covert retrieval may affect working memory and long-term learning was consistent with studies by Loaiza, McCabe, Youngblood, Rose and Myerson (2011) who found evidence for levels of processing effects in immediate complex span tasks and later tests of delayed free recall. In a reading span task, participants were asked to perform a series of sentence verification tasks whilst remembering the terminal words for a later test of serial order. Deep sentences prompted participants to respond to the semantic characteristics of the to-be-remembered word, such as “The brother of one of your parents is an UNCLE”; “A tool for making clothes is a sewing MACHINE”. Shallow sentences prompted participants to respond to the orthographic characteristics of the to-be-remembered word, such as “A word made up of five letters is UNCLE”; “There are three different vowels in the world MACHINE”. Loaiza et al. found a recall advantage for deep processing in both the reading span tasks and later tests of delayed free recall. Similarly, when participants were presented with four concrete words that were interleaved with arithmetic sums, the authors found a recall advantage for conditions in which participants made a deep, semantic judgement (was the word living/non-living?) relative to a shallow, orthographic judgement (does the word have more than two vowels?) in both the operation span tasks and later tests of delayed free recall.

Rose, Buchsbaum and Craik (2014) also showed that covert retrieval could be used in a test of working memory when both rehearsal and refreshing were disrupted. They presented participants with five blocks of 24 trials in which single words were presented for either a deep (living/nonliving) or shallow (“e”/no “e”) levels of processing judgment. Following the judgement, participants engaged in a 10s retention interval in which they were (a) encouraged to rehearse, (b) performed easy mental arithmetic (and so were assumed to use attentional refreshing but not rehearsal) or (c) performed hard mental arithmetic (and so were assumed to be able to use neither rehearsal nor attentional refreshing and had to rely on covert retrieval). They then recalled the word. At the end of the experiment, there was a 10-minute filler task and then a surprising final free recall test. In the test after 10s, performance in the rehearsal condition was fast and close to ceiling, there was no need for covert retrieval from episodic memory and there were no levels of processing effect. By contrast, recall after 10s of easy and hard math was slower and less accurate and there were levels of processing effects suggesting covert retrieval. In the surprise final free recall task, there were levels of processing effects for all three conditions, and now the two math-filler conditions outperformed the rehearsal group. Rose et al. argue that a covert retrieval mechanism may have refreshed the stimuli during the easy math filler task, and was necessary to account for successful immediate recall in the two filler conditions. These covert retrievals were beneficial in final test of free recall, as these words now outperformed the rehearsal conditions that did not require covert retrieval.

Loaiza and McCabe (2012) found further support for the covert retrieval explanation of the McCabe (2008) effect. They again showed that subspan ISR was superior to complex span in an immediate test; but in a delayed test, items presented for complex span were recalled better than those presented for subspan ISR. They also included a supra-span test of ISR. This was included to provide an additional control for the extra difficulty in the initial list. With this control, accuracy in the immediate recall of the first four items of the 8-item list was now equivalent to that in complex span, but nevertheless in a delayed test, items presented for complex span were still recalled far better than those presented for ISR, as predicted if it were the opportunities for covert retrieval within complex spans tasks that generated the McCabe effect.

Later work (e.g., Loaiza and McCabe, 2013; Loaiza, Duperreault, Rhodes & McCabe, 2015) argued that the covert retrieval interpretation of the McCabe effect was also consistent with an attentional refreshing account, but in an attempt to discriminate between the covert retrieval and the attentional refreshing accounts of the McCabe effect, Loaiza and Halse (2019) examined the effects of list length and distractors in immediate and delayed tests. Across three experiments, participants were first presented with lists of 2-4 words, then completed 0-3 arithmetic problems, and then performed serial recall. Loaiza and Halse (2019) argued that increasing list length would reduce participants' opportunities to perform attentional refreshing, but increasing the number of distractors at the end of the list would increase the requirement and opportunities to use covert retrieval. They also argued that early list items would receive more opportunities for refreshing than those presented at later serial positions. They found that with an immediate test, serial recall declined with increasing list length, serial recall declined with increasing distractors, and there were primacy effects with serial recall scoring. However, in a delayed test of free recall, there was no effect of list length, no effect of serial position, but recall increased with increasing distractors. Loaiza and Halse (2019) interpreted these findings as evidence for a covert retrieval rather than the attentional refreshing account.

However, a covert retrieval interpretation has been questioned by Souza and Oberauer (2017) who argued that it was the additional time taken to encode the items (rather than the covert retrieval) that resulted in superior delayed recall of complex span over simple span items. Souza and Oberauer replicated the original McCabe effect using the original methodology in which only simple and complex span tasks were used. However, they performed additional experiments in which they required participants to perform simple span trials, complex span trials, and very slow simple span tasks (in which words were presented for ISR at very slow rates equivalent to the complex span tasks). When all three trial types were tested together, there was no McCabe effect: simple span and slow span were better recalled than complex span items in an immediate test, but items presented for complex span were no better recalled in a delayed test than those presented for simple spans (indeed, delayed recall was best for the ISR items that had been presented at very slow rates). In an echo of classic rehearsal-based interpretations, the authors argue that delayed recall performance was related to the time spent during encoding.

Two recent studies have examined whether participants use any additional time at study to perform elaborative (interactive imagery) rehearsal in working memory tasks, and if so, have examined whether this would lead to benefits in immediate tests and long-term retention. Bartsch, Singmann and Oberauer (2018) presented participants with lists of six words to remember in serial order. The words appeared at a fast rate of 0.5s/item in boxes arranged vertically from the top to bottom of the screen. After the last word had been presented, the first or second triplet of words underwent further processing: the words were either repeated or their locations were cued for attentional refreshing and this was performed with or without interactive vivid imagery instructions. Bartsch et al. found no benefit of either elaboration or refreshing on an immediate test of recognition, but there was a benefit from elaboration instructions but not refreshing in a later test of recognition in which participants were presented with stimuli and asked to recognise other members of the probed triplet.

Thalmann, Souza and Oberauer (2019, Experiment 3) also showed a benefit of long-term elaboration on a final delayed test of recognition, and also showed limited benefits of elaboration with concrete words in delayed serial recall. Thalmann et al. presented participants with short lists of four abstract or concrete words that they were required to maintain for 10s prior to serial recall. One group was instructed to perform elaboration, the other received no such instructions. In counterbalanced blocks, participants performed the task with or without articulatory suppression (continuously saying "babibu"). All participants also performed a surprise delayed recognition task. When performance on the 4-word lists was first tested after 10s, recall was better for the concrete words than for the abstract words, and recall was better in the no suppression conditions relative to the articulatory suppression conditions. Within the suppression conditions, there was a recall advantage for the elaboration instructions that was limited to the concrete words. In a delayed test of recognition, there were again advantages for the concrete words over the abstract words, but there were now recognition benefits for words that had been encoded under articulatory suppression, and there was an effect of elaboration on delayed recognition for words that had not been encoded under articulatory suppression.

Finally, a recent review by Hartshorne and Makovski (2019) suggests that there is after all, an overall positive effect of even maintenance rehearsal on long-term learning. Hartshorne and Makovski presented evidence from 13 new large-scale experiments and conducted a meta-analysis on these and 61 prior experiments. Both prior and new experiments showed evidence of improvement in a delayed test for items that have been maintained longer in working memory. Importantly, the meta-analysis suggests that the effect of maintenance rehearsal on long-term retention is similar for tests of recall and tests of recognition (cf. Greene, 1987). The meta-analysis also found no evidence that later long-term retention was affected by whether or not there was an original test during the working memory phase.

The Relationship between Rehearsal and Repetitions

In Bjork's (1970) early article, he not only addressed the role of rehearsal and forgetting in the Brown-Peterson task, but he also considered the representation of repetitions and the representation of rehearsals. In particular, he questioned whether a successful rehearsal of an item operates in essentially the same manner as a presentation of that item?

Considering first repetitions, Bjork argued that repetitions generally improved episodic memory and he considered three possible mechanisms for this: a repetition might increase the cumulative strength or activation of a single trace of the repeated item, a repetition might create an additional memory trace of the repeated item, or something else (e.g., a repetition might lead to increasing transfer from STS to LTS)? Bjork described the task of discriminating between these options as "formidable" and that the lack of relevant data sets at that time made the task even more difficult.

More than 50 years later, there is no lack of empirical data examining the effects of repetition in episodic memory (e.g., Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Delaney, Verkoeijen, & Spirgil, 2010) and it is generally agreed that (i) episodic memory is better for spaced or distributed repetitions compared with massed repetitions (e.g., Greene, 1992; Kahana & Howard, 2005), that (ii) there is a generally positive relationship between increased recall and increased lag between repetitions (Madigan, 1969), and that (iii) the optimal lag between repetitions is proportional to the retention interval (e.g., Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Glenberg, 1976).

But how are items that are repeated or rehearsed represented? The idea that repetitions serve only to increase the cumulative strength of an item appeared to have been ruled out by participants' related abilities to make two independent judgements of serial position to twice-presented stimuli and participants' ability to make judgements of frequency and judgements of recency (e.g., Flexser & Bower, 1974; Hintzman & Block, 1971, for a review, see Hintzman, 2010). A multiple trace theory appears better placed to explain how participants can make these different judgements to different repetitions.

One major theory that proposes multiple traces is the contextual variability theory, according to which, each presented item is associated with a slowly drifting contextual representation (e.g., Estes, 1955; Melton, 1970). Repeated items can benefit from having multiple routes by which contextual representations can be used to cue the target item, and the lag effect arises because the benefit of a repetition on later recall is increased to the extent that the context associated with the repetition is distinct from the context associated with the original presentation.

If each repetition resulted in an independent memory trace, then the recall probability of a repeated stimulus could be estimated based on the recall of two separate once-presented items. However, Benjamin and Tullis (2010) argue that contextual variability theory cannot account for superadditivity effects and non-monotonicity in lag functions. Rather, they argue that during the study of a repeated item, the participants may be reminded that they have already studied that item, such that the memory traces are not always independent but may be interactive and interdependent. This reminding (or study-phase retrieval) becomes less likely as the lag between the first and second presentation increases, but when it occurs, the act of retrieval potentiates memory in a way that is positively related to the difficulty of retrieval.

Consistent with a multiple traces view of repetition, Hintzman (2010) found that participants could make multiple recency judgements for items that there were presented three times. However, he also found that participants' later recency judgements were affected by earlier recency judgements made to the same items. Hintzman argued that this violation of independence was caused by recursive reminding. During the second presentation of a word, the participants were reminded of the first presentation and used this to make a recency judgement. This experience of being reminded was encoded into memory. During the third presentation, they were reminded of the earlier reminding experience resulting in recursive reminding, and the incorporation of the first inter-item interval information affected the estimate of the second inter-item interval.

The ideas of recursive reminding are appealing as they capture the phenomena of everyday experiences (Hintzman, 2011) and they help explain the effects of repetition of identical and related stimuli in the laboratory (e.g., Tullis, Benjamin & Ross, 2014), especially when remembering the temporal order of events is important. For example, Wahlheim and Jacoby (2013) presented participants with two lists of paired associates that included pairs that were different (A-B, C-D), pairs that were repeated on the two lists (A-B, A-B), and pairs had the same stimulus term but a changed response term (A-B, A-D) on the two lists. During the presentation of the second list, participants were given the opportunity to say whether they thought that the response term of a pair had changed in List 2, and if they could remember, to say what the response term had been in List 1. At test, participants were presented with the stimulus terms from the second list and asked to recall the most recent response term. They were also asked to report whether another response came to mind prior to or simultaneously with their final response. Consistent with a recursive reminding (or memory for change) account, Wahlheim and Jacoby (2013) found that performance on the A-D pairs reflected a mixture of proactive interference and proactive facilitation. When a change was not detected, the earlier A-B learning led to proactive interference on later A-D learning relative to the control condition. By contrast, when a change was detected and later recollected, recall of A-D was actually facilitated relative to the control condition. The role of recursive reminding and memory for change has also been extended to both proactive and retroactive interference (Jacoby, Wahlheim & Kelley, 2015), recency judgements (Jacoby & Wahlheim, 2013), and has been used to explain findings for spaced repetitions in cued recall (Wahlheim, Maddox & Jacoby, 2014).

Let us consider now the similarities and differences between rehearsals and repetitions. Like repetitions, later recall benefits when stimuli are rehearsed more frequently (Rundus, 1971; Tan & Ward, 2000), more recently (Brodie & Murdock, 1977; Tan & Ward, 2000) and when the distribution of the rehearsals is more spaced (Modigliani & Hedges, 1987; Tan & Ward, 2000). There is also clear evidence that rehearsal can be recursive or cumulative (e.g., Laming, 2006) with future rehearsal sequences building on earlier rehearsals sequences.

That participants experience reminding in rehearsals as well as repetitions has been demonstrated by Rundus (1971, Experiment 4). Rundus presented participants with lists of 24 words for free recall using the overt rehearsal methodology. The lists were composed of 12 unrelated items and six exemplars each from two semantic categories, and they were presented in a random order. Rundus found that during the presentation of the study list, participants' choices of which words to rehearse were highly influenced by the category of the just-presented item. There was a far higher tendency to rehearse a word that was just-previously rehearsed if it was from the same category as the just-presented item (74%) than from a different category (16%) or unrelated item (21%). There was also a far higher tendency to rehearse a word from the same category (61%) than from a different category (6%) or unrelated item (8%) even if the word hadn't been just-previously rehearsed. Interestingly, the unrelated items formed something of their own subjective category. When an unrelated item was presented at study, there was a far higher tendency to rehearse another unrelated word that was just-previously rehearsed (42%) than a just-previously rehearsed item from a different category (21%). There was also a greater tendency to rehearse other unrelated items that had not just-previously been rehearsed (15%) than other category exemplars (7%).

Further evidence of reminding using overt rehearsals was provided by McKinley and Benjamin (2020). In their two experiments, participants were presented with pairs of related and unrelated words that were separated by various lags. They found that participants were much more likely to rehearse related words rather than unrelated words in the interval immediately following a presented list item. Moreover, these remindings predicted enhanced later recognition (Experiment 1) and cued recall (Experiment 2).

Finally, the relationship between rehearsals and repetitions is strengthened by studies that have directly examined the mnemonic consequences of repeating and rehearsing items. Murdock and Metcalfe (1978) presented participants with 32 lists of 20 words at a slow presentation rate (1 word every 5s) for IFR using the overt rehearsal method. The words were presented for 1s, and during the 4s of blank screen that followed, participants were instructed to continually report aloud whatever they were thinking of during the inter-presentation intervals. There were no restrictions on which words they should think about or the method by which they should try to encode them. These overt rehearsal lists not only provided recall data but also allowed the authors to determine the mean distribution of the words rehearsed after each presented word and the lag distribution of the rehearsed words for each participant. The authors used the distributions of rehearsals to generate controlled rehearsal schedules that were tailored for each participant. In a further 32 controlled-rehearsal lists, participants saw a study word and then were presented with computer-generated "rehearsals" based on the rehearsal statistics of that participant. Murdock and Metcalfe (1978) showed striking similarities between the serial position curves from the overt rehearsal conditions and the serial position curves from the controlled-rehearsal conditions, and there were very similar functional serial position curves between the two conditions, even when the data were further partitioned by the frequency of rehearsals.

Tan and Ward (2000) also compared recall following overt rehearsal with recall following presented 'rehearsals'. In one experiment, participants were presented with lists of 20 high frequency (HF) and lists of 20 low frequency (LF) words for free recall using the overt rehearsal method. There was a HF recall advantage in the early nominal serial positions and a HF advantage throughout the functional serial position curve. The early HF words were also rehearsed more frequently than the early LF words. To examine these word frequency effects, a later experiment presented three groups of participants with words for study and 'rehearsal'. The participants of the yoked-HF group received the content and were presented with the rehearsals of HF participants from the earlier experiment, the yoked-LF group received the content and rehearsals of LF participants from the earlier experiment, and the yoked-LF-with-HF-rehearsals group received the LF study words of the LF participants but using rehearsal schedules generated by the HF group. Tan and Ward found that the serial position curves of the yoked-HF group did not differ significantly from the serial position curves of the HF group, the serial position curves of the yoked-LF group did not differ significantly from the serial position curves of the LF group, and recall in the yoked-LF-with-HF-rehearsals group was at an intermediate level between the yoked-HF and the yoked-LF groups. This suggests that repetitions had similar mnemonic benefit to rehearsals on later recall, and that word frequency effects are partly mediated by rehearsal schedules.

In summary, these studies suggest that there are striking similarities between the effects of repetitions and rehearsals on recall. Recall benefits from increases in the frequency and spacing of both repetitions and rehearsals, and a repeated or related stimulus can remind the participant of an earlier presented or rehearsed item. Moreover, when rehearsals are compared directly with presented repetitions there are striking similarities in the resulting serial position curves. These results may seem surprising given that recall advantages are normally associated with retrieval practice and testing relative to restudy (e.g., Roediger & Karpicke, 2006), but one should remember that the studies reported here often used near-immediate tests, and the testing effects are often more pronounced following an extended delay.

Summary of the Current Points of Contention and a Personal Viewpoint

Despite 50 or more years of research on rehearsal, there remain fundamental differences of opinion as to the effectiveness of rehearsal as a memory strategy. Across different short-term memory or working memory tasks, we have seen that: (i) rehearsal has been proposed to *maintain* items in serial order in memory, thereby offsetting the effects of trace decay (e.g., Baddeley, 1986; Barrouillet, Bernardin, & Camos, 2004), (ii) rehearsal has been proposed by others to *enhance* later accessibility by strengthening associations in LTS (Atkinson & Shiffrin, 1968, 1971; Raaijmakers & Shiffrin, 1981) or creating multiple copies of the stimuli that vary in their frequency, recency, and distribution of rehearsal (Tan &

Ward, 2000), and (iii) rehearsal has been proposed to have *no causal role* in support of serial recall (Lewandowsky & Oberauer, 2015; Souza & Oberauer, 2018, 2020).

Nevertheless, the data reviewed here point toward a positive relationship between rehearsal and recall in the Brown-Peterson task (Hellyer, 1962; Ward et al., 2019), in IFR (Rundus, 1971; Tan & Ward, 2000), and (when participants are free to rehearse how they like) in ISR (Tan & Ward, 2008) and in the complex span task (Ward et al., 2019). Regarding serial recall tasks, participants rarely choose to rehearse the entire list when they have complete freedom to rehearse how they wish, consistent with the claims that participants benefit from rehearsing a subspan set of items (Jarrold, 2017), through to the end of the list (Barrouillet et al., in press). Moreover, when participants are encouraged to rehearse longer sequences, this results in no overall benefit in ISR and complex span (Souza & Oberauer, 2018, 2020), presumably because inaccuracies in rehearsal sequences occur at longer list lengths. We have also found evidence supporting a reasonably close relationship between rehearsals and repetitions in long-term episodic memory together with renewed interest in the mnemonic consequences of rehearsal (whether this shows benefits of covert retrieval or additional study time) in immediate tasks on later retention.

A Personal take on rehearsal processes

I would like to use the final subsection to present a personal take on the material reviewed in this chapter. To my mind, the review of repetition effects showed evidence consistent with the claim that repeated items are represented by multiple inter-dependent memory traces. We saw that a later test of memory benefits from increased numbers of spaced repetitions, we saw that participants were able to make multiple recency and frequency judgements to repeated items, and we saw that the traces were not entirely independent: participants may be reminded of earlier events during the presentation of related or repeated items (e.g., Benjamin & Tullis, 2010; Hintzman, 2011; Wahlheim & Jacoby, 2013). Such findings are consistent with theories of memory in which a representation of each repetition of an item is associated with a slowly drifting temporal context. Upon encoding, the item itself may act as a retrieval cue (study-phase retrieval) and the retrieved information may be rehearsed (McKinley & Benjamin, 2000; Rundus, 1971), encoded, and used to further update the temporal context (e.g., Howard & Kahana, 2002; Lohnas, Polyn & Kahana, 2015; Polyn, Norman, & Kahana, 2009; Siegel & Kahana, 2014). The review also showed that there were striking similarities between the mnemonic benefits of different patterns of rehearsals and corresponding patterns of repetitions on later recall.

The idea that items are associated with a time-varying context can help explain serial position effects in the IFR of long lists of items. After the last list item has been presented, the time-varying context at test can be used as a cue. Later list items are more likely to be associated with contexts that are more similar to the context at the end of the list, and this similarity gives rise to extended recency effects. Moreover, the temporal context of one item will be similar to that of its neighbours, resulting in temporal contiguity effects at output (Howard & Kahana, 1999; Kahana, 1996). Within such accounts, rehearsing an item increases the number and variety of associated contexts, generating more retrieval routes, and rehearsing an item to later functional serial positions will increase its accessibility because it will become associated with contexts that are more similar to the end-of-list-context used at test (Tan & Ward, 2000). Moreover, a recalled word may help recall others: words are more likely to be output consecutively that have been presented consecutively (Howard & Kahana, 1999; Kahana, 1996), last co-rehearsed to similar temporal contexts (Ward et al., 2003), and recently rehearsed in the same sequential order (e.g., Laming, 2006, 2008).

Lists presented for ISR are similarly encoded to lists presented for IFR (Bhatarah et al., 2008; Grenfell-Essam & Ward, 2012). For the shorter list lengths (more typical of ISR), participants tend to initiate recall of short lists of words in IFR and ISR with the first item and continue by recalling in a forwards order (Ward et al., 2010; Ward & Tan, 2019). The exact reason for this is uncertain, and requires further investigation, but a variety of mechanisms in IFR have been proposed including: a start-of-list context cue (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Farrell, 2012; Metcalfe & Murdock, 1981), the increased temporal distinctiveness of the first item (e.g., Brown, Neath, & Chater, 2007), increased attention (e.g., Lohnas, Polyn, & Kahana, 2015), or a “Get Ready” warning signal (e.g., Laming, 1999, 2010). In ISR, possible mechanisms include: that the first item may be encoded with the greatest strength (e.g., Page & Norris, 1998), that the first item may be associated with a start-list cue (e.g., Farrell, 2012; Henson, 1998), or that the first item may be associated with early context positions (e.g., Burgess & Hitch, 1992, 1999, 2006). Recall sequences that start with the first list item contribute greatly to the extended primacy effects in IFR and this tendency is of course essential for accurate ISR performance (Ward et al., 2010).

At fast presentation rates, ISR (Tan & Ward, 2008) and IFR (Grenfell-Essam et al., 2013) of short lists can be performed reasonably well, even in the absence of rehearsal. However, recall benefits from a slower rate in both tasks when participants have greater opportunities to rehearse: there is a higher tendency to initiate recall with the first list item and higher overall accuracy. If the rate is slowed under articulatory suppression, the tendency to initiate recall with the first item declines very steeply with increasing list length in IFR (Grenfell-Essam et al., 2013) leading to a decrease in IFR under articulatory suppression at slow rates. In ISR, there are some studies that suggest serial recall declines at slower rates under articulatory suppression (Baddeley & Lewis, 1984) whereas others have shown an increase in serial recall at slower rates, even under articulatory suppression (Longoni, Richardson & Aiello, 1993) and further work is necessary to understand these differences.

My own preference is to assume that the processes underpinning rehearsal are the same as the processes underpinning retrieval (Laming, 2006) and it is clear that the output from retrieval (the rehearsals and remindings) are themselves encoded, and one can examine the development of the rehearsal sequences and relate these to sequences of recalls (e.g., Laming, 2006, 2008). There are a number of advantages to this approach. First, it is self-evident that we can retrieve, and no one questions whether retrieval serves an important memory function. Second, if rehearsals and recalls are underpinned by the same processes, then rehearsal is likely to enhance later recall to the extent that an earlier retrieval is likely to enhance a later retrieval: this might be more likely following spaced retrieval attempts but less likely following massed retrieval attempts. Third, such an approach suggests that to the extent that we can retrieve information, we can rehearse information, therefore providing a parsimonious approach to dealing with the retention and recall of verbal and non-verbal stimuli (e.g., Cortis, Dent, Kennett, & Ward, 2015; Cortis Mack, Dent & Ward, 2018; Jones, Farrand, Stuart, & Morris, 1995). In addition, this approach constrains the variety of different retrieval strategies available at rehearsal to the different retrieval strategies available at test, and may set limits on the upper bounds of when serial rehearsal remains effective (prior to when serial retrieval becomes error-prone). Finally, all computational models of recall have a retrieval mechanism, but not all have a computational account of rehearsal (e.g., Brown, Neath & Chater, 2007; Farrell, 2012; Howard & Kahana, 2002), and it may be provocative to consider whether the former can simply serve as the latter (cf., Burgess & Hitch, 1999) or whether different equations dictate rehearsal and recall (e.g., Raaijmakers & Shiffrin, 1981).

Summary and Conclusions

This chapter has reviewed rehearsal processes in human memory, including the different proposed functions of rehearsal, the different methods used in its study, and the different roles of rehearsal in different memory tasks. We have found that in free recall, rehearsal tends to improve later memory, especially if the rehearsals are distributed and recent. We have found that the effectiveness of rehearsal on later serial recall is more contentious, but rehearsal tends to improve later serial recall if a subspan sequence of early items can be rehearsed towards the end of the list. To my mind, there are sufficient similarities between the patterns of rehearsal and recall across the different tasks reviewed here to encourage a unified account, and my preference would be to liken the act of rehearsal to the act of retrieval where the consequences of rehearsal on later retention are commensurate with the mnemonic consequences of equivalently distributed repetitions.

References

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable?. *Child development*, *77*(6), 1698-1716.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of Learning and Motivation (Vol. 2*, pp. 89-195). Academic Press.

Atkinson, R.C., & Shiffrin, R.M. (1971). The control of short-term memory. *Scientific American, 225*, 82-90.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5(3),* 119-126.

Awh, E., Jonides, J., & Reuter-Lorenz, P. (1998). Rehearsal in Spatial Working Memory. *Journal of Experimental Psychology: Human Perception and Performance, 24(3),* 780-790.

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. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.) *Recent advances in learning and motivation, Vol. 8.* (pp. 47-90). London: Academic Press.

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. Cortex, 112, 91-106.

Baddeley, A., & Lewis, V. (1984). When does rapid presentation enhance digit span?. *Bulletin of the Psychonomic Society, 22(5),* 403-405.

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. Quarterly *Journal of Experimental Psychology, 36A*, 233-252.

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

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource-sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*, 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and Cognitive Load in Working Memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 33(3)*, 570-585.

Barrouillet, P., & Camos, V. (2012). As time goes by: temporal constraints in working memory. *Current Directions in Psychological Science, 21,* 413–419. doi: 10.1177/0963721412459513

Barrouillet, P., & Camos, V. (2015). *Working Memory: Loss and Reconstruction*. Hove, UK: Psychology Press.

Barrouillet, P., Gorin, S., & Camos, V. (in press). Simple Spans Underestimate Verbal Working Memory Capacity. *Journal of Experimental Psychology: General*

Barrouillet, P., Portrat, S., Vergauwe, E., Diependaele, K., & Camos, V. (2011). Further Evidence for Temporal Decay in Working Memory: Reply to Lewandowsky and

Oberauer (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1302-1317.
Bartsch, L. M., Singmann, H., & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory & Cognition*, *46*(5), 796-808.
Bayliss, D. M., Bogdanovs, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language, 81*, 34-50.
Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective?. *Cognitive Psychology, 61(3),* 228-247.
Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition, 37*, 689–713. https://doi.org/10.3758/MC.37.5.689
Bhatarah, P., Ward, G., & Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: The serial nature of recall and the effect of test expectancy. *Memory & Cognition, 36*, 20-34.
Bjork, R. A. (1970). Repetition and rehearsal mechanisms in models for short-term memory. *Models of human memory*, 307-330.
Brodie, D. A. (1975). Free recall measures of short-term store: Are rehearsal and order of recall data necessary?. *Memory & Cognition*, *3*(6), 653-662.
Brodie, D. A., & Murdock Jr, B. B. (1977). Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior, 16,* 185-200.
Brodie, D. A., & Prytulak, L. S. (1975). Free recall curves: Nothing but rehearsing some items more or recalling them sooner?. *Journal of Verbal Learning and Verbal Behavior*, *14*(5), 549-563.
Brown J. (1958). Some tests of the decay theory of immediate memory. *The Quarterly Journal of Experimental Psychology, 10,* 12–21. https://doi.org/10.1080/17470215808416249
Brown, G. D., Della Sala, S., Foster, J. K., & Vousden, J. I. (2007). Amnesia, rehearsal, and temporal distinctiveness models of recall. *Psychonomic Bulletin & Review*, *14*(2), 256-260.
Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed?. *Journal of Memory and Language, 34(5)*, 594-621.
Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539-576.
Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31(4),* 429-460.

Burgess, N., & Hitch, G.J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.
Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55(4*), 627-652.
Camos, V. (2015). Storing verbal information in working memory. Current Directions in Psychological *Science, 24*, 440–445.
Camos, V. (2017). Domain-specific vs. domain-general maintenance in working memory: Reconciliation within the time-based resource sharing model. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 67, pp. 135–171).

Camos, V., & Barrouillet, P. N. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: the executive and phonological loops. *Frontiers in Human Neuroscience, 8*, 900.

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory?. *Annals of the New York Academy of Sciences*, *1424*(1), 19-32.

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language, 61*, 457– 469.

Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition, 39*, 231–244. https://doi.org/10.3758/s13421-010-0011-x

Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *The Quarterly Journal of Experimental Psychology Section A*, *45*(2), 177-192.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis. *Psychological Bulletin, 132(3),* 354-380.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19(11),* 1095-1102.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769-786.

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1179-1214.

Cortis Mack, C., Dent, K., & Ward, G. (2018). Near-independent capacities and highly constrained output orders in the simultaneous free recall of auditory-verbal and visuo-spatial stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 107-134.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*, 163-191.

Cowan, N. (1995, 2016). *Working Memory Capacity: Classic Edition*. Routledge.

Cowan, N. (1999) An embedded-processes model of working memory. In *Models of Working Memory* (Miyake, A. and Shah, P., Eds.), pp. 62–101, Cambridge University Press.

Craik, F. I. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, *9*(2), 143-148.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 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*(3), 268-294.

Craik, F. I., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12(6),* 599-607.

Davelaar, E., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H., & Usher, M. (2005). The Demise of Short-Term Memory Revisited: Empirical and Computational Investigations of Recency Effects. *Psychological Review, 112(1),* 3-42.

De Schrijver, S., & Barrouillet, P. (2017). Consolidation and restoration of memory traces in working memory. *Psychonomic Bulletin & Review, 24,* 1651-1657.

Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, *54*(3), 180-187.

Delaney, P. F., Verkoeijen, P. P., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In *Psychology of Learning and Motivation (Vol. 53,* pp. 63-147). Academic Press.

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. The Quarterly Journal of Experimental Psychology, 60(9), 1227-1245.

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71*, 43–51.

Estes, W. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62(3)*, 145-154.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory for letter strings. Journal of Verbal Learning and Verbal Behavior, 12(4), 360-372.

Farrell, S. (2012). Temporal clustering and sequencing in short- term memory and episodic memory. *Psychological Review, 119,* 223–271. doi:10.1037/a0027371

Fischler, I., Rundus, D., & Atkinson, R. C. (1970). Effects of overt rehearsal procedures on free recall. *Psychonomic Science*, *19*(4), 249-250.

Flexser, A. J., & Bower, G. H. (1974). How frequency affects recency judgments: a model for recency discrimination. *Journal of Experimental Psychology, 103(4),* 706-716.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39,* 3–27. doi:10.1111/1469- 7610.00301

Gathercole, S. E. & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory & Language, 28*, 200-213.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology*, *40*(2), 177-190.

Glanzer, M., & Meinzer, A. (1967). The effects of intralist activity on free recall. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 928-935.

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 1-16.

Glenberg, A., Smith, S. M., & Green, C. (1977). Type I rehearsal: Maintenance and more. *Journal of Verbal Learning and Verbal Behavior*, *16*(3), 339-352.

Greene, R. L. (1987). Effects of Maintenance Rehearsal on Human Memory. *Psychological Bulletin*, *102*(3), 403-413.

Greene, R. L. (1992). *Human Memory: Paradigms and Paradoxes*. Psychology Press.

Grenfell-Essam, R., & Ward, G. (2012). Examining the relationship between free recall and immediate serial recall: The role of list length, strategy use, and test expectancy. *Journal of Memory and Language, 67*, 106-148.

Grenfell-Essam, R., & Ward, G., & Tan, L. (2013). The role of rehearsal on the output order of immediate free recall of short and long lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 317-347.

Guitard, D., Gabel, A. J., Saint-Aubin, J., Surprenant, A. M., & Neath, I. (2018). Word length, set size, and lexical factors: Re-examining what causes the word length effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1824.

Hartshorne, J. K., & Makovski, T. (2019). The effect of working memory maintenance on long-term memory. *Memory & Cognition*,47, 749-763.

Healy, A. F., & McNamara, D. S. (1996). Verbal learning and memory: Does the modal model still work? *Annual Review of Psychology, 47*, 143-172.

Hellyer, S. (1962). Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology, 64*, 650. http://dx.doi.org/10.1037/h0043117

Henry, L. A., Messer, D., Luger-Klein, S., & Crane, L. (2012). Phonological, visual, and semantic coding strategies and chil- dren's short-term picture memory span. *The Quarterly Journal of Experimental Psychology, 65*, 2033–2053. doi:10.1080/ 17470218.2012.672997

Henson, R. N. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36(2),* 73-137.

Hintzman, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, *32*(2), 336-350.

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition, 38(1),* 102-115.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*(3), 253-271.

Hintzman, D. L., & Block, R. A. (1971). Repetition and Memory: Evidence for a Multiple-Trace Hypothesis. *Journal of Experimental Psychology, 88(3),* 297-306.

Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition, 3,* 197–209.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 923-941.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46,* 269-299.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38(2),* 241-253.

Jacoby, L. L. (1973). Encoding processes, rehearsal, and recall requirements. *Journal of Verbal Learning and Verbal Behavior*, *12*(3), 302-310.

Jacoby, L. L., & Wahlheim, C. N. (2013). On the importance of looking back: The role of recursive remindings in recency judgments and cued recall. *Memory & Cognition*, *41*(5), 625-637.

Jacoby, L. L., Wahlheim, C. N., & Kelley, C. M. (2015). Memory Consequences of Looking Back to Notice Change: Retroactive and Proactive Facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41(5),* 1282-1297.

Jahnke, J. C. (1965). Primacy and recency effects in serial-position curves of immediate recall. *Journal of Experimental Psychology*, *70*(1), 130-132.

Jalbert, A., Neath, I., Bireta, T. J., & Surprenant, A. M. (2011). When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 338-353.

Jarrold, C. (2017). The Mid-Career Award: Working out how working memory works: Evidence from typical and atypical development. *The Quarterly Journal of Experimental Psychology, 70(9*), 1747-1767.

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2010). The nature and position of processing determines why forgetting occurs in working memory tasks. *Psychonomic Bulletin & Review, 17(6),* 772-777.

Johnson, R. E. (1980). Memory-based rehearsal. In Psychology of Learning and *Motivation (Vol. 14*, pp. 263-307). Academic Press.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21(4),* 1008-1018.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24,* 103-109.

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review, 12(1)*, 159-164.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189– 217

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. Journal of Verbal learning and Verbal Behavior, 1, 153-161.

Kroll, N. E., & Kellicutt, M. H. (1972). Short-term recall as a function of covert rehearsal and of intervening task. Journal of Verbal Learning and Verbal Behavior, 11(2), 196-204.

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology, 34,* 419-426.

Laming, D. (2006). Predicting free recalls. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32,* 1146-1163.

Laming, D. (2008). An improved algorithm for predicting free recalls. *Cognitive Psychology, 57,* 179–219.

Laming, D. (2009). Failure to recall. *Psychological Review, 116,* 157–186.

Laming, D. (2010). Serial position curves in free recall. *Psychological Review, 117,* 93-113.

Langerock, N., Vergauwe, E., & Barrouillet, P. (2014). The maintenance of cross-domain associations in the episodic buffer. *Journal of Experimental Psychology: Learning, Memory & Cognition, 40,* 1096-1109.

Lehman, M., & Malmberg, K. J. (2013). A Buffer Model of Memory Encoding and Temporal Correlations in Retrieval. *Psychological Review, 120(1),* 155-189.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, *11*(5), 771-790.

Lewandowsky, S., Nimmo, L. M., & Brown, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language, 58*, 415-428.

Lewandowsky, S., & Oberauer, K. (2015). Rehearsal in serial recall: An unworkable solution to the non-existent problem of decay. *Psychological Review, 122*, 674-699. http://dx.doi.org/10.1037/a0039684

Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences, 13(3)*, 120-126.

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review, 122(2),* 337-363.

Loaiza, V. M., Duperreault, K. A., Rhodes, M. G., & McCabe, D. P. (2015). Long-term semantic representations moderate the effect of attentional refreshing on episodic memory. *Psychonomic Bulletin & Review, 22(1)*, 274–280. https://doi.org/10.3758/s13423-014-0673-7

Loaiza, V., & Halse, S. (2019). Where Working Memory Meets Long-Term Memory: The Interplay of List Length and Distractors on Memory Performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(8),* 1455-1472.

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(2)*, 191–203. https://doi.org/10.3758/s13421-011-0148-2

Loaiza, V. M., & McCabe, D. P. (2013). The influence of aging on attentional refreshing and articulatory rehearsal during working memory on later episodic memory performance. *Aging, Neuropsychology, and Cognition, 20(4),* 471–493.

Loaiza, V. M., McCabe, D. P., Youngblood, J. L., 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*, 1258-1263.

Longoni, A. M., Richardson, J. T., & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition, 21(1),* 11-22.

Lovatt, P., Avons, S. E., & Masterson, J. (2000). The word-length effect and disyllabic words. *The Quarterly Journal of Experimental Psychology Section A*, *53*(1), 1-22.

Madigan, S. A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior, 8(6),* 828-835.

McFarlane, K. A., & Humphreys, M. S. (2012). Maintenance rehearsal: the key to the role attention plays in storage and forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1001-1018.

McKinley, G. L., & Benjamin, A. S. (2020). The role of retrieval during study: Evidence of reminding from overt rehearsal. *Journal of Memory and Language, 114,* 104128.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior, 9(5),* 596-606.

Metcalfe, J., & Murdock, B. B. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning and Verbal Behavior, 20(2),* 161-189.

Meunier, G., Ritz, D., & Meunier, J. (1972). Rehearsal of individual items in short-term memory. *Journal of Experimental Psychology, 95(2),* 465-467.

Meunier, G., Stanners, R., & Meunier, J. (1971). Pronounceability, rehearsal time, and the primacy effect of free recall. *Journal of Experimental Psychology, 88(1),* 123-127.

Modigliani, V., & Hedges, D. G. (1987). Distributed Rehearsals and the Primacy Effect in Single-trial Free Recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 426-436.

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64,* 482-488.

Murdock, B., & Metcalfe, J. (1978). Controlled rehearsal in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *17*(3), 309-324.

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*(2), 480-494.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology, 21*, 263–276.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679–684.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*(1), 53-81.

Naveh-Benjamin, M., & Ayres, T. J. (1986). Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology, 38(A),* 739–751.

Neath, I., Bireta, T. J., & Surprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review, 10*, 430-434.

Neath, I., & Brown, G. D. A. (2006). Further applications of a local distinctiveness model of memory. *Psychology of Learning and Motivation, 46,* 201–243.

Neath, I., & Nairne, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review*, *2*(4), 429-441.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short term and working memory. Psychological Bulletin, 144(9), 885-958.

Oberauer, K. (2019). Is Rehearsal an Effective Maintenance Strategy for Working Memory? Trends in Cognitive Sciences, 23(9), 798-809.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in Immediate Serial Recall: Decay, Temporal Distinctiveness, or Interference? *Psychological Review*, *115*(3), 544-576.

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the Time-Based Resource- Sharing theory. *Psychonomic Bulletin & Review, 18*, 10–45. http://dx.doi.org/10.3758/s13423-010-0020-6

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H.-M. (2005). Working memory and intelligence—Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 61–65

Page, M., & Norris, D. (1998). The Primacy Model: A New Model of Immediate Serial Recall. *Psychological Review, 105(4),* 761-781.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193-198.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116(1),* 129-156.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*, 93–134. http://dx.doi.org/10.1037/0033-295X.88.2.93

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex*, *43*(1), 135-145.

Reitman, J. S. (1974). Without surreptitious rehearsal, information in short-term memory decay. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 365-377.

Ricker, T. J., & Cowan, N. (2014). Differences between Presentation Methods in Working Memory Procedures: A Matter of Working Memory Consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(2),* 417-428.

Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: insights from studies on attention and working memory. *Annals of the New York Academy of Sciences, 1424(1),* 8-18. https://doi.org/10.1111/nyas.13633

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *The Quarterly Journal of Experimental Psychology*, *69*(10), 1969-1995.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17(3),* 249-255.

Rose, N. S., Buchsbaum, B. R., & Craik, F. I. M. (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(5)*, 689–700. https://doi.org/10.3758/s13421-014-0398-x

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89*, 63–77. http://dx.doi.org/10.1037/h0031185
Rundus, D. (1977). Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(6), 665-681.
Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior, 9*, 99-105.
Ryan, J. (1969). Grouping and short-term memory: Different means and patterns of grouping. *The Quarterly Journal of Experimental Psychology*, *21*(2), 137-147.
Schweikert, R. & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 12*, 419-425.
Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40(3),* 755-764.
Souza, A. S., & Oberauer, K. (2020). No Evidence That Articulatory Rehearsal Improves Complex Span Performance. *Journal of Cognition, 3(1),* 11-11.
Standing, L., Bond, B., Smith, P., & Isely, C. (1980). Is the immediate memory span determined by subvocalisation rate? *British Journal of Psychology, 71*, 525-539.
Shaffer, W., & Shiffrin, R. M. (1972). Rehearsal and storage of visual information. *Journal of Experimental Psychology*, *92*(2), 292-296.
Souza, A. S., & Oberauer, K. (2017). Time to process information in working memory improves episodic memory. *Journal of Memory and Language*, *96*, 155-167.
Souza, A. S., & Oberauer, K. (2018). Does articulatory rehearsal help immediate serial recall? Cognitive Psychology, 107, 1-21. http://doi.org/10.1016/j.cogpsych.2018.09.002
Spurgeon, J., Ward, G., & Matthews, W. J. (2014). Examining the relationship between immediate serial recall and immediate free recall: common effects of phonological loop variables but only limited evidence for the phonological loop. *Journal of Experimental Psychology. Learning, Memory, & Cognition, 40*, 1110-1141. http://dx.doi.org/10.1037/a0035784
Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1589–1625. http://dx.doi.org/10.1037/0278-7393.26.6.1589
Tan, L. & Ward, G. (2007). Output order in immediate serial recall. *Memory & Cognition, 35*, 1093–1106.
Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review, 15*, 535–542. http://dx.doi.org/10.3758/PBR.15.3.535
Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(1),* 37-55.
Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory & Cognition, 28*, 341-348.
Tullis, J. G., Benjamin, A. S., & Ross, B. H. (2014). The reminding effect: Presentation of associates enhances memory for related words in a list. *Journal of Experimental Psychology: General*, *143*(4), 1526-1540.
Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, *49*(4), 446-468.
Unsworth, N., & Engle, R. W. (2007). 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. http://dx.doi.org/10.1037/0033-295X.114.1.104

Vallar, G., & Baddeley, A. D. (1982). Short-term forgetting and the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, *34*(1), 53-60.

Vergauwe, E., & Langerock, N. (2017). Attentional refreshing of information in working memory: Increased immediate accessibility of just-refreshed representations. *Journal of Memory and Language*, *96*, 23-35.

Vergauwe, E., Langerock, N., & Barrouillet, P. (2014). Maintaining information in visual working memory: memory for bindings and memory for features are equally disrupted by increased attentional demands. *Canadian Journal of Experimental Psychology, 68,* 158-162. doi:http://doi.org/10.1037/cep0000025

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive effects of memory. *Memory & Cognition, 41(1),* 1-15.

Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The Role of Reminding in the Effects of Spaced Repetitions on Cued Recall: Sufficient but Not Necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(1),* 94-105.

Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory & Cognition, 30*, 885-892.

Ward, G., Cortis Mack, C., Doherty, C., Knight, N. & Loaiza, V. (2019, November 14-17). *Rehearsal Improves Recall in Short-Term and Working Memory*. [Paper Presentation] Psychonomic Society 60th Annual Meeting, Montreal, Quebec, Canada.

Ward, G., & Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: A reexamination using overt rehearsal. *Journal of Experimental Psychology: Learning Memory and Cognition, 30*, 1196-1210.

Ward, G. & Tan, L., (2019). Control processes in short-term storage: Retrieval strategies in immediate recall depend upon the number of words to be recalled. *Memory & Cognition, 47, 658-682.* http://dx.doi.org/10.3758/s13421-018-0891-8

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1207–1241. http://dx.doi.org/10.1037/a0020122

Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 29*, 186-210.

Watkins, M. J., Peynircioĝlu, Z. F., & Brems, D. J. (1984). Pictorial rehearsal. *Memory & Cognition*, *12*(6), 553-557.

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

Wickelgren, W. A. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, *68*(4), 413-419

Wickelgren, W. A. (1967). Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology*, *19*(2), 97-102

Woodward Jr., A. E., Bjork, R. A., & Jongeward Jr, R. H. (1973). Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 608-617.

Figure Captions

Figure 1. Brown-Peterson data from Ward et al. (2019, Experiment 2). The mean recall accuracy as a function of the number of computer-controlled rehearsals and the retention interval using Brown-Peterson scoring (Panel A), Serial Recall scoring (Panel B), and Free Recall scoring (Panel C).

Figure 2. Immediate Free recall (IFR) data from Ward (2002) using the overt rehearsal method. Panel A shows the mean proportion of words recalled as a function of the nominal serial position. Panel B shows the mean frequency of rehearsals afforded to individual words as a function of the nominal serial position. Panel C shows the mean proportion of words recalled plotted by the functional serial position (Panel C), which is the rank-order in which the items were most recently rehearsed. Panel D shows the mean proportion of words recalled plotted by when each item was most recently rehearsed. Reprinted with permission from Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory & Cognition, 30*, 885-892.

Figure 3. Immediate Serial Recall (ISR) data of 6-word lists from Tan and Ward (2008) using the overt rehearsal method. Panel A shows the serial position curves for words presented at slow, medium and fast presentation rates; Panel B shows the rehearsal strategies at slow rates; and Panel C shows the relationship between ISR accuracy and the maximum rehearsed sequence length at slow rates. Note: Given the small sample size in this study, the effect size estimate in C should be treated with caution. Edited and reprinted with permission from three figures from Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review, 15*, 535–542. http://dx.doi.org/10.3758/PBR.15.3.535

Figure 4. Complex span data from Ward et al. (2019, Experiment 1) using the overt rehearsal method. Panel A shows the serial position curves for lists of between 1 and 6 words; Panel B shows the rehearsal strategies for the 6-item lists; and Panel C shows the relationship between ISR accuracy and the maximum rehearsed sequence length for the 6-item lists. Note: Given the small sample size in this study, the effect size estimate in C should be treated with caution.

Figure 1

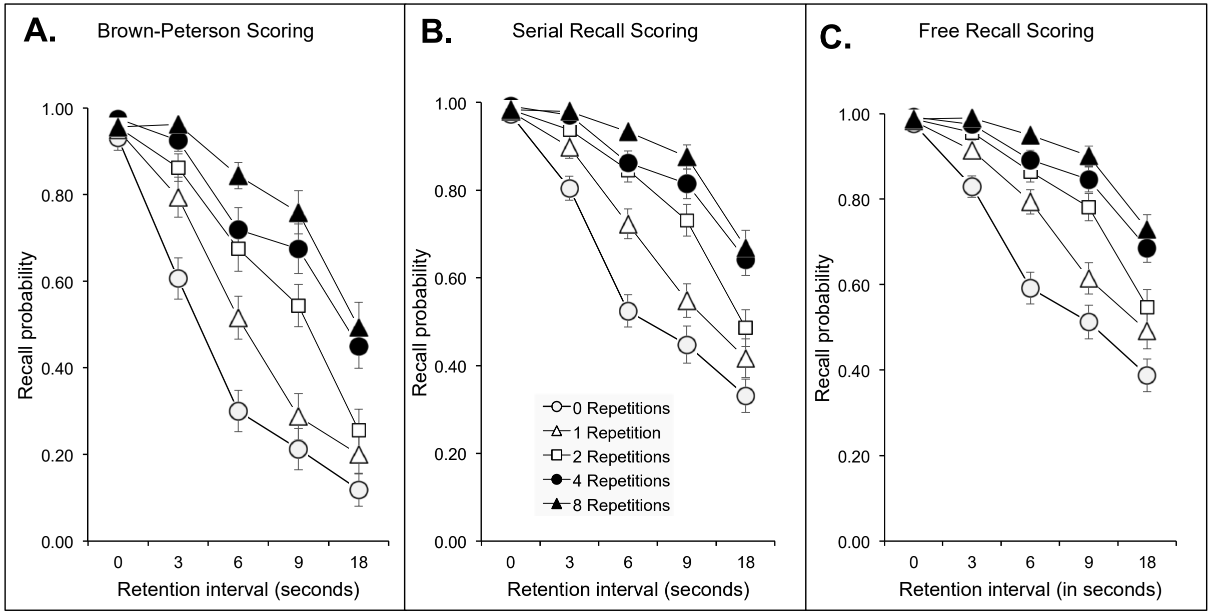

A.
Brown-Peterson Scoring
B.
Serial Recall Scoring
C.
Free Recall Scoring
Recall probability
1.00
0.80
0.60
0.40
0.20
0.00
0
3
6
9
18
Retention interval (seconds)
Retention interval (in seconds)
0 Repetitions
1 Repetition
2 Repetitions
4 Repetitions
8 Repetitions

Figure 2

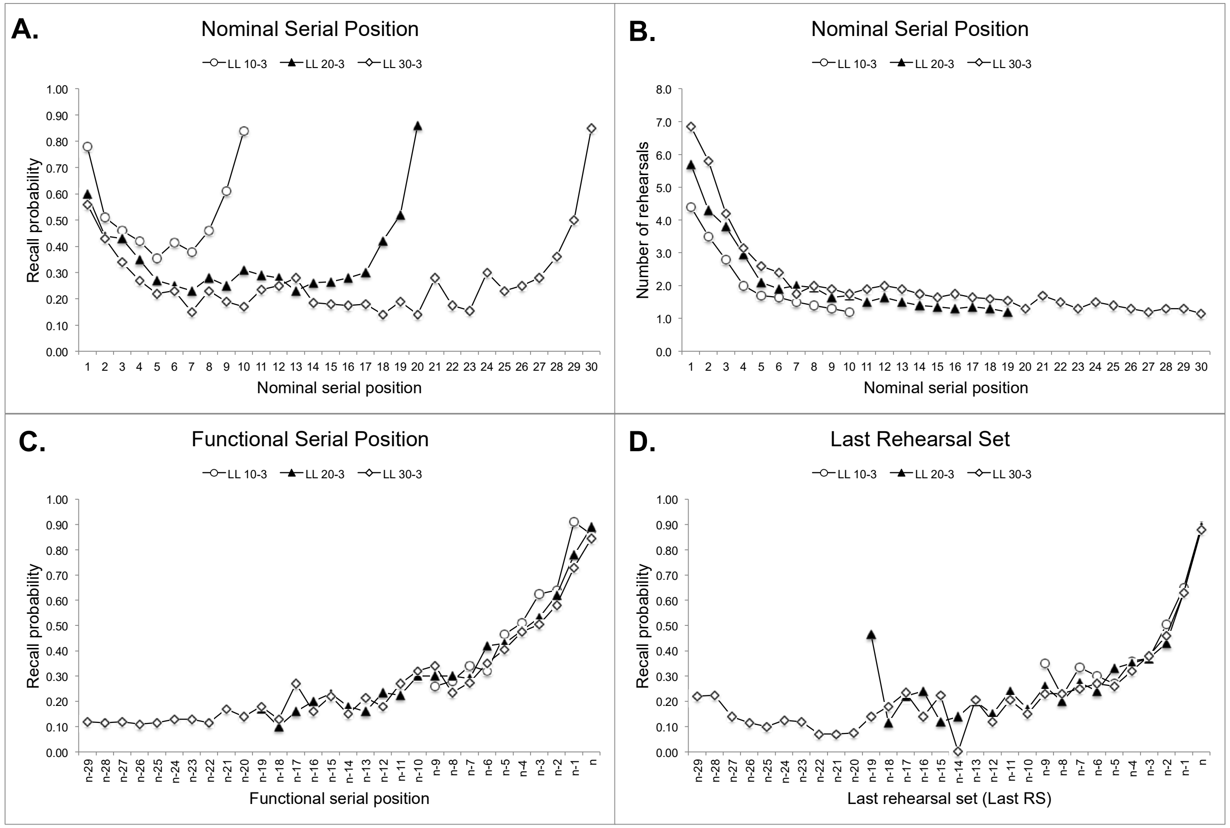

A.
Nominal Serial Position
LL 10-3
LL 20-3
LL 30-3
Recall probability
Nominal serial position
B.
Nominal Serial Position
LL 10-3
LL 20-3
LL 30-3
Number of rehearsals
Nominal serial position
C.
Functional Serial Position
LL 10-3
LL 20-3
LL 30-3
Recall probability
Functional serial position
D.
Last Rehearsal Set
LL 10-3
LL 20-3
LL 30-3
Recall probability
Last rehearsal set (Last RS)

Figure 3

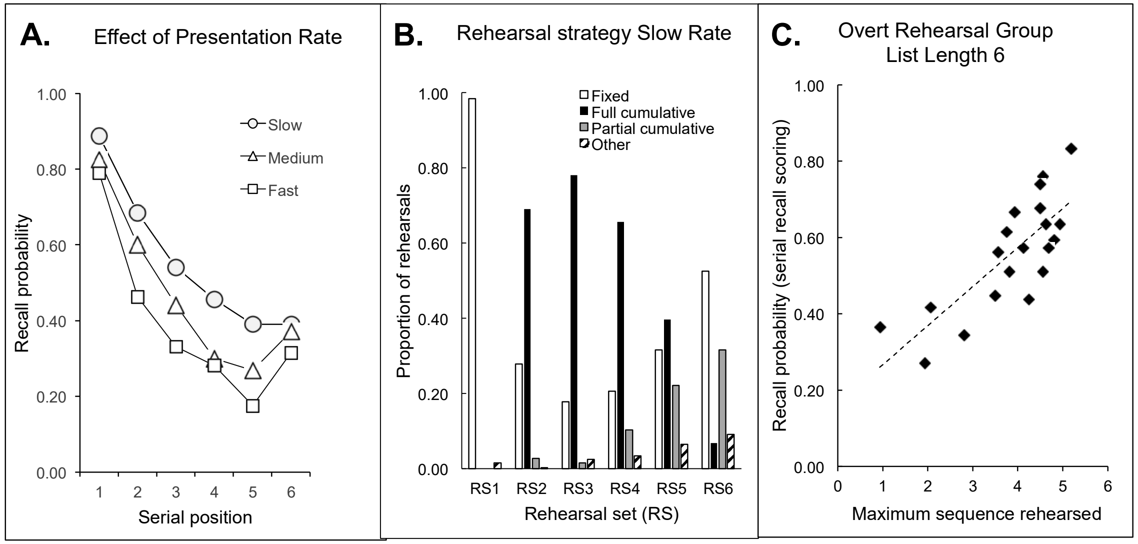
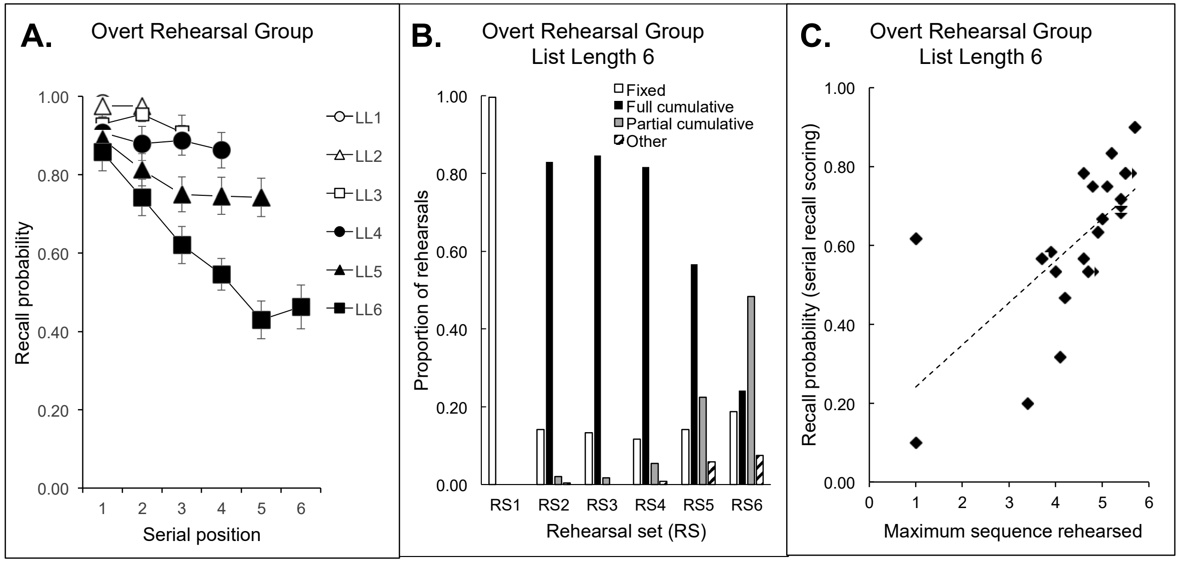

A.
Effect of Presentation Rate
Slow
Medium
Fast
Recall probability
Serial position
B.
Rehearsal strategy Slow Rate
Fixed
Full cumulative
Partial cumulative
Other
Proportion of rehearsals
RS1 RS2 RS3 RS4 RS5 RS6
Rehearsal set (RS)
C.
Overt Rehearsal Group
List Length 6
Recall probability (serial recall scoring)
Maximum sequence rehearsed

Figure 4


A.
Overt Rehearsal Group
Recall probability
Serial position
LL1
LL2
LL3
LL4
LL5
LL6
B.
Overt Rehearsal Group
List Length 6
Proportion of rehearsals
Rehearsal set (RS)
Fixed
Full cumulative
Partial cumulative
Other
RS1
RS2
RS3
RS4
RS5
RS6
C.
Overt Rehearsal Group
List Length 6
Recall probability (serial recall scoring)
Maximum sequence rehearsed