Running Head: PROCESSING IN WORKING MEMORY

How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects

Christopher Jarrold[1], Helen Tam[1], Alan D. Baddeley[2], and Caroline E. Harvey[1]

[1]University of Bristol

[2]University of York

Abstract

Three studies are presented that examine why the processing demands within working memory tasks lead to forgetting of the memoranda. In each, separate groups of adult participants were asked to carry out either verbal or nonverbal operations on exactly the same processing materials, while maintaining verbal storage items. The imposition of verbal processing tended to produce greater forgetting, despite the fact that verbal processing operations took no longer to complete than nonverbal processing operations. However, nonverbal processing did cause forgetting, relative to baseline control conditions, and evidence from the timing of individuals' processing responses suggested that individuals in both processing groups slowed their responses in order to 'refresh' the memoranda. Taken together the data suggest that processing has a domain-general effect on working memory performance by impeding refreshment of memoranda, but can also cause effects, which appear domain-specific, either by blocking rehearsal or as a result of interference. In addition, the balance of these effects depends on the structure of the working memory task employed.

How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects

Complex span tasks are seen by many to be the 'gold standard' measure of working memory (Conway, et al., 2005; Cowan et al., 2005; Jarrold & Towse, 2006). In complex span tasks, participants are asked to encode lists of memoranda (which are to be recalled later) while intermittently performing some kind of distracting processing activity (Case, Kurland, & Goldberg, 1982; Daneman & Carpenter, 1980). In general, complex span performance is inferior to that of simple span, which measures immediate recall of memoranda after their presentation in the absence of distracting processing activity (e.g., Duff & Logie, 2001; Hutton & Towse, 2001; La Pointe & Engle, 1990). Consequently, and perhaps unsurprisingly, the presence of the processing activity within a complex span task tends to cause forgetting of memoranda presented during the task.

One reason for the substantial research interest in the complex span task is that performance on this measure is typically a strong correlate of measures of intelligence in adults (Bayliss, Jarrold, Gunn, & Baddeley, 2003; Kyllonen & Christal, 1990; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002) and of academic achievement in children (Bayliss et al., 2003; Bull, Epsy, & Wiebe, 2008; Swanson, 2008). Furthermore, many would argue that complex span performance is a significantly stronger correlate of these abilities than is simple span (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Luaghlin, & Conway, 1999; Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & Süß, 2005), indicating that the forgetting caused by the imposition of processing within a complex span task increases its predictive power. What is less clear is why processing causes forgetting in working memory paradigms, and this is a question we seek to address in this work. At the same time, however, a number of studies exist in which complex and simple span tasks produce similar levels of performance and show comparable correlations with aptitude measures (see Ackerman, Beier, & Boyle, 2005; Bayliss, Jarrold, Baddeley, & Gunn, 2005; Colom, Rebello, Abad, & Shih, 2006; Cowan et al., 2005), raising a second question of when the inclusion of processing within a working memory paradigm does or does not lead to substantial and meaningful forgetting.

Broadly speaking there are two possible explanations of the role of processing within working memory paradigms: domain-general or domain-specific accounts (Guérard & Tremblay, 2008). One of the most influential current accounts of working memory function is provided by the time-based resource-sharing (TBRS) model of Barrouillet, Camos, and colleagues (Barrouillet, Bernadin, & Camos, 2004; Barrouillet, Bernadin, Portrat, Vergauwe, & Camos, 2007; Camos, Langer, & Barrouillet, 2009). This model postulates that attention is shared between two activities that occur within the processing interval of a complex span task – the processing requirement itself, and the covert maintenance of to-be-remembered items already presented on the trial. A key method of covert maintenance assumed by the model is a domain-general attention-dependent process whereby memory representations are reactivated or "refreshed" (cf. Hudjetz & Oberauer, 2007). The amount of attention required to perform processing – attention that is required to retrieve the result of a processing operation from long-term memory and which could have otherwise been used for refreshment – is indexed by "cognitive load", a measure based on the difficulty and the rate of the processing demand. Importantly, this conceptualisation of cognitive load implies that regardless of the nature of the processing demand, so long as it is capable of capturing attention by requiring retrieval from long-term memory, refreshment will be compromised. Indeed, as Vergauwe, Barrouillet, and Camos (2009) claim, "the central attentional resource has to be time-shared between processing and storage regardless of the nature and domain of the information involved" (p. 1013). To illustrate this point, Vergauwe, Barrouillet, and Camos (in press) gave participants complex span tasks in which they either had to encode letters (verbal storage) or spatial positions (visuo-spatial storage) in between performing either verbal or visuo-spatial processing judgements. They found that both types of processing led to comparable forgetting of material, regardless of whether memoranda were verbal or visuo-spatial. This, and related findings from this group (see Barrouillet et al., 2007, Experiment 3; Vergauwe et al., 2009) has led to the assumption within the TBRS model that refreshment operations depend entirely on domain-general attentional processes (see also Morey & Cowan, 2005).

However, as Vergauwe et al. (in press) themselves note (see also Camos et al., 2009), this evidence of domain-general forgetting due to processing is somewhat surprising given the large number of studies elsewhere which suggest domain-specificity, rather than domain-generality, in the effect of processing on recall. From the outset, the literature on working memory has been replete with evidence that the imposition of a dual-task processing load leads to domain-specific disruption of memory (Baddeley & Hitch, 1974). One of the most striking demonstrations of this effect was provided by Hale, Myerson, Rhee, Weiss, Abrams (1996) who asked participants to remember either sequences of digits or spatial locations while either carrying out a verbal (saying the colour of each stimulus as it appeared) or visual (identifying the colour of each stimulus by pointing to a response grid) secondary task. Relative to control conditions with no dual task requirements, digit recall was selectively impaired by colour naming while recall of spatial locations was selectively impaired by pointing (see also Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002; Farmer, Berman, & Fletcher, 1986; Guérard & Tremblay 2008; Logie, Zucco, & Baddeley, 1990; Smyth & Pelky, 1992). Similar evidence of specific effects of processing on storage has also been observed in complex span tasks. For example, Shah and Miyake (1996) crossed type of processing (verbal or spatial) with type of storage (verbal or spatial) to create four different versions of the complex span task. They found that span was greater in tasks where there was a mismatch between storage and processing type than in tasks where storage and processing materials were from the same domain or stimulus class (see also Bayliss, et al., 2003; Conlin, Gathercole, & Adams, 2005, Maehara & Saito, 2007; Turner & Engle, 1989). Indeed, when Vergauwe et al. (in press) examined the linear relationships between cognitive load and recall in their study they found that increasing the cognitive load of verbal processing had a greater detrimental effect on the recall of verbal memoranda than did corresponding increases to the cognitive load of visuo-spatial processing. Consequently, while the Vergauwe et al. (in press) study does provide evidence for domain-general effects of processing, it also contains evidence of domain-specific causes of forgetting.

One prominent explanation of domain-specific patterns of forgetting is that different maintenance operations are blocked by the imposition of different dual tasks demands. For example, Baddeley (1986) argued that verbal material is maintained in working memory via phonological rehearsal, a maintenance activity that is selectively disrupted by verbal dual tasks (Peterson & Johnson, 1971; Saito, Logie, Morita, & Law, 2008). Similarly, Logie (1995) argued for a corresponding visuo-spatial 'rehearsal' system that is selectively blocked by visuo-spatial dual task activity (Smyth & Pelky, 1992). An alternative view is that the processing of dual task stimuli that are similar to the to-be-remembered information causes interference that may alter, or "overwrite" the features of the representations stored in memory, such that the originals can no longer be retrieved (Lewandowsky, Geiger, & Oberauer, 2008; Lewandowsky, Oberauer, & Brown, 2009; Oberauer, 2009; Oberauer & Lewandowsky, 2008).[1] Of course, these two explanations of apparently domain-specific patterns of forgetting are not necessarily mutually exclusive; interference may occur precisely when the on-line maintenance of memoranda is prevented by a particular dual task (cf. Unsworth, Heitz, & Parks, 2008). In addition, one could potentially argue that working memory performance is supported by both domain-general and domain-specific processes. In a recent clarification of the TBRS model Camos et al. (2009) have argued that domain-general refreshment and domain-specific rehearsal operations might both operate to support recall in complex span tasks, and might both be potentially affected by the processing demands on this task (see also Vergauwe et al., in press).

In this paper we attempt to reconcile, to some degree at least, the apparently confusing and contradictory data in the literature on whether the processing demands of working memory tasks cause forgetting via domain-general or domain-specific mechanisms. First, we note that while a number of studies have systematically manipulated different classes of storage and processing operations within the complex span procedure (e.g., Bayliss et al., 2003; Conlin et al., 2005; Shah & Miyake, 1996), these have typically presented different stimuli for, for example, either verbal or visual processing. This raises the possibility that the cognitive load of these different processing operations differs, which in turn might account for apparent domain-specific forgetting effects, even under a domain-general account such as the TBRS model (Vergauwe et al., 2009). The one exception to this is the Maehara and Saito (2007) study in which different storage loads were systematically paired with the same processing requirement. However, although Maehara and Saito manipulated the domain of processing employed in their complex span procedure *across* experiments, they did not directly compare the effects of different domains of processing of a known cognitive load within the same study. In the three studies presented below we make this form of comparison, which allows us to directly address the issue of whether the nature (domain-specific account) or the difficulty (domain-general account) of processing leads to forgetting from working memory.

In addition, another potential inconsistency in the previous literature, highlighted above (see also Vergauwe et al., 2009), is that some studies of forgetting from working memory have employed the complex span procedure while others have utilised dual task approaches in which memoranda and processing are presented concurrently (Farmer et al., 1986; Logie et al., 1990; Hale et al., 1996; Guérard & Tremblay, 2008) or when a pre-load of to-be-remembered items is followed by a single block of processing (Cocchini et al., 2002; Klauer & Zhao, 2004; Smyth & Pelky, 1992). There are two, related, differences between complex span and dual task approaches that might have important implications for how interpolated processing gives rise to forgetting. First, the fact that participants switch between successive processing and storage operations in a complex span task means that participants might strategically delay their attempt to carry out the processing operation that follows a given storage item in order to refresh, rehearse, or consolidate that item, or the list of items presented up to that point in the task (Barrouillet et al., 2004; Lépine, Bernadin, & Barrouillet, 2005; Towse, Hitch, & Hutton, 2002). Second, in complex span tasks storage items are presented incrementally, so that the kind of maintenance or consolidation operations that can take place during these switches between processing and storage can benefit, at least at the start of the trial, from being focussed on only a subset of the memoranda. For example, after presentation of the first item participants only have to refresh, rehearse or consolidate that item, and so such operations are likely to be effective. In contrast in dual tasks situations when all the memoranda are presented in a block, any refreshment, rehearsal or consolidation is more likely to operate on the whole list, and consequently is less likely to be successful. Taken together, this implies that there may be more opportunities for 'micro switches' between processing and storage in complex span tasks, which might serve to refresh memoranda or consolidate items into long-term memory. Conversely, in dual task paradigms there may be a greater premium placed on trying to maintain the whole list of items via more strategic and effortful rehearsal activity.

To test these suggestions the current studies contrasted two types of task (cf. Tehan, Hendry, & Kocinski, 2001). One was the standard complex span task in which storage and processing operations were inter-leaved. The other was a version of a pre-load dual task, which we term a Brown-Peterson task because of parallels with this classic procedure (Brown, 1958; Peterson & Peterson, 1959), in which a comparable number of storage items and the same total duration of processing was presented, but with these phases of the task blocked. Because of the concern about the comparability of different types of processing used in previous studies, in our studies a single set of stimuli served as the processing stimuli in these two tasks, but the type of processing (verbal or nonverbal) to be performed on these items was manipulated between-participants through pre-experimental instructions. In order to control for any underlying differences in storage capacity of the two groups of participants employed in each study, and to directly examine the degree of forgetting caused by the imposition of a processing load in the Brown-Peterson task, a final task was included (see Figure 1). This was a delayed span task, which was identical to the Brown-Peterson task except that the single block of processing was replaced by an unfilled delay.

The memoranda used in all task conditions were verbal. Consequently, in the complex span and Brown-Peterson tasks, half of the participants carried out same-domain (verbal) processing whereas the other half carried out different-domain (nonverbal) processing. Because both types of processing involved the retrieval of knowledge from long-term memory, (i.e., neither was a mere simple reaction response), it can be safely assumed that both types of processing captured attention, and hence should disrupt domain-general attentional refreshment (Barrouillet et al., 2004, 2007). However, the two types of processing were designed to ensure that phonological rehearsal would be more disrupted under verbal than nonverbal processing, and clearly the likelihood of similarity-based interference occurring is greater when the memoranda are paired with verbal than with nonverbal processing. In addition, in each study efforts were made to ensure that the difficulty, or cognitive load, associated with the verbal processing task was no greater than that associated with the nonverbal processing task. As a result, any evidence of greater forgetting due to the requirement of having to perform verbal as opposed to nonverbal operations on the common processing stimuli can be taken as strong evidence of domain-specific forgetting.

Experiment 1

In Experiment 1, participants were presented with a complex span, Brown-Peterson, and delayed span tasks in which the to-be-remembered items were single syllable words. The processing operations included in the complex span and Brown-Peterson tasks involved the presentation of a circular array of small circles. However, half of the sample (verbal processing group) was required to judge whether the total number of small circles shown was odd or even. As the number of items presented was always beyond the readily subitisable range (Dehaene, 1992), we reasoned that the counting involved in making this decision would require verbal processing which would potentially block rehearsal. Participants in the other half of the sample (nonverbal processing group) saw exactly the same displays but were required to judge whether a circle that differed from the others, and which was always present, appeared on the left or the right of the display. As no verbal reasoning was required in making this decision, we assumed that subvocal rehearsal would be possible during processing for individuals in this group.

Method

Participants

The participants were recruited from the University of Bristol. All were native speakers of English and between 17 and 35 years of age. They either received course credits or £7 in return for their participation. Data from two participants who produced unacceptably high error rates (over 25%) in their processing judgments were discarded. In the end, data from 30 participants (7 males) were available for analysis, with 15 participants in each of the verbal (same-domain) and nonverbal (different-domain) processing groups.

Tasks

All participants were given three memory span tasks – complex span, Brown-Peterson span, and delayed span – as well as two baseline processing tasks, one given before and one after the span tasks. As is shown schematically in Figure 1, at a particular span length, trials in all three memory tasks contained the same number of storage items, and were of the same total duration. In the complex span task, each storage item was immediately followed by an interval of the processing requirement. The intervals were equal in duration and, in total, the duration of all intervals equalled to that of the single block of processing which followed the presentation of the item list in the Brown-Peterson task. In the delayed span task, storage items were presented in succession, followed by an unfilled delay before recall was prompted.

Materials

Ninety-two monosyllabic concrete nouns were used as storage items in each of the three tasks. Each item was presented both auditorily and pictorially. The items were read by a male voice, and each was equated to 750 ms in duration by the addition of silence at the end of the recording. The pictorial forms of the items were black-and-white line drawings, all sized to fit within a square of approximately 7cm by 7cm, adapted from Snodgrass and Vanderwart (1980) with additional drawings created for those not provided by the Snodgrass and Vanderwart set.

The storage items were used to create four trials at each of the span lengths (4, 5, 6, and 7 items), and an additional practice trial containing four storage items. The memoranda in the practice trial were identical in all three tasks. Across the three tasks, the same set of items appeared at a particular span length, but these items were presented in different combinations such that all trials were different across tasks.

The processing requirement (in the two baseline tasks, and the Brown-Peterson and complex span tasks) consisted of either location or parity judgments on a series of arrays made up of circles. All circles within the array were of the same size (diameter = 13 mm), and each had a small gap (2 mm) in their circumferences, with the exception of one circle that had two of these small gaps. Participants in the nonverbal processing group were to decide whether the circle with two gaps in it was located on the left or right side of the screen, whereas participants in the verbal processing group were simply instructed to decide if the total number of circles in the array was odd or even. There were equal numbers of arrays constructed that contained 5, 6, 7, or 8 circles. In each array, each circle was positioned at one of twelve predetermined locations that were all equidistant from the central point of the screen, with the target circle for individuals in the nonverbal processing group being situated on either side of the screen an equal number of times.

Procedure

Participants were tested individually using a laptop. The session began with the first baseline processing task. Participants in the nonverbal processing group were shown examples of the “target” and “distractor” circles, and were instructed to find the target in each circle array. They were told to indicate the side of the screen the target circle was on by pressing the corresponding key on the keyboard – “q” for left, and “p” for right. Participants in the verbal processing group were not alerted to the difference between target and distractor circles, but were simply asked to count the number of circles in the array, and then decide if the total was odd or even by pressing “q” for odd or “p” for even. After these instructions, four practice trials (with feedback on response accuracy) were given. These practice trials were followed by 30 test trials, where no feedback was given. All trials began with the presentation of a fixation cross at the centre of the screen for 250 ms, followed by a blank screen of 25 ms, which was in turn followed by the circles array. The array remained on the screen until a response was made, and the next trial followed after an inter-trial-interval of 25 ms. At test, participants were not given feedback on their accuracy. The responses and response times (RTs) on all trials were recorded by the computer.

After the first baseline task, participants were then given the three memory span tasks – complex span, Brown-Peterson, and delayed span. The order of presentation of these tasks was counterbalanced across participants using a Latin square design. For each task, participants were told that they would see a series of items being presented, and they were to remember these items for recall later. In the complex span task, participants were informed that each item in the series would be followed by a brief period of circle array judgments, and depending on their group membership, they were to perform either the location or parity judgments on these arrays. In the Brown-Peterson task, participants were told that the items would be presented one after the other, and the item list would be followed by the circle arrays. In the delayed task, participants were instructed that after the items had been shown, there would be a period of delay before recall was required. In all three tasks, recall was prompted at the end of the trial by the appearance of a question mark ("?") on the screen. Participants were asked to verbally recall the objects in the same order as they were presented, and to say "don't know" in place of the ones that they had forgotten.

All three memory span tasks began with a practice trial of four storage items. This was followed by the test proper, starting with trials of a span length of four storage items and progressing (with an increment of one storage item at a time) up to seven storage items. There were four test trials at each span length level. On each trial, the storage items were presented auditorially, and also pictorially at the center of the screen for 750 ms. An inter-stimulus-interval of 250 ms followed each storage item. In the processing intervals within the Brown-Peterson and complex span tasks, the circle arrays were presented in the same manner as in the baseline task. In the complex span task, the duration of each processing episode was 6 s. The duration of the processing interval on a Brown-Peterson trial was therefore 24 s, 30 s, 36 s, and 42 s for span lengths of 4, 5, 6, and 7 items respectively. The delay periods on the delayed span trials were of the same durations for the corresponding span length. On all tasks, the recall prompt ("?") appeared at the centre of the screen at the end of each trial, and the participant's item recall was recorded by hand by the experimenter.

After all three memory span tasks had been administered, participants were given the second baseline task to end the session. The procedure for this task was identical to the first baseline task, except that there were no practice trials. The experimental session in its entirety lasted one hour.

Results

Table 1 presents summary statistics for each processing group for speed and accuracy of baseline processing carried out pre- (baseline 1) and post- (baseline 2) test, and for memory performance. Baseline processing RTs were calculated by averaging the median RTs obtained from each participant's correct responses only.

Baseline Processing Tasks

RT. A 2 (task: pre-test/ post-test) x 2 (processing group: verbal/nonverbal) mixed ANOVA was performed on the median RT obtained from the pre- and post-test baseline tasks. This analysis showed a significant task effect, $F(1, 28) = 66.16$, $p < .001$, MSE = 15269.75, _^{2} = .703, as responding was faster in the baseline task given after than before the memory tasks. Participants in the verbal processing group had numerical lower RTs than their counterparts in the nonverbal processing group (see Table 1), but this difference was not significant, $F(1, 28) < 1$, $p = .429$. The task x processing group interaction was also not statistically significant, $F(1, 28) = 3.16$, $p = .086$.

Accuracy. A 2 (task: pre-test/ post-test) x 2 (processing group: verbal/nonverbal) mixed ANOVA on proportion of correct responses for the pre- and post-test baseline tasks revealed a significant main effect of group, $F(1, 28) = 9.47$, $p = .005$, $MSE = .001$, _^{2} = .253. Both processing groups achieved a high level of accuracy on the baseline measures, but the nonverbal processing group was on average more accurate than the verbal group. Neither the task main effect nor the task x processing group interaction was significant, $F(1, 28) < 1$, $p = .411$ and $F(1, 28) = 2.26$, $p = .144$ respectively.

Memory Span Tasks

Recall Scores. In each memory span task, the proportion of items recalled for each trial was first determined for each participant, and the proportions recalled were summed across trials to establish the recall score (see Table 1). As there were four trials at each of the four span length levels (4, 5,6 and 7 items), scores could range from 0 to 16.

A 3 (task: complex span/ Brown-Peterson/ delayed span) x 2 (processing group: verbal/nonverbal) mixed ANOVA was performed on the recall scores. There was a significant effect of task, $F(2, 56) = 32.23$, $p < .001$, $MSE = 0.964$, _^{2} = .535. Post-hoc comparisons showed that this significant task effect arose because recall was better in the complex span than delayed span task, $t(29) = 4.67$, $p < .001$. In turn, better recall was achieved in the delayed span than in the Brown-Peterson task, $t(29) = 3.84$, $p = .001$.

In contrast, the main effect of processing group was not significant, $F(1, 28) = 1.02$, $p = .32$, _^{2} = .035. The task x processing group interaction, however, approached statistical significance, $F(2, 56) = 2.91$, $p = .063$, $MSE = 0.964$, _^{2} = .094. This interaction was explored in two ways. First, the effect of task was examined in each group separately. Among individuals in the verbal processing group, there was a suggestion of superior performance on the complex span task than on the delayed span task, $t(14) = 2.08$, $p = .056$, and performance on the delayed span task was significantly higher than that on the Brown-Peterson task, $t(14) = 3.74$, $p = .002$. Among individuals in the nonverbal processing group, complex span performance was clearly superior to delayed span scores, $t(14) = 4.88$, $p < .001$, but delayed span scores were not significantly higher than Brown-Peterson scores, $t(14) = 1.73$, $p = .105$. Second, difference scores relative to delayed span were obtained for each individual's performance on both complex span and Brown-Peterson span. This was done because the delayed span task contained no processing, and consequently was identical in form for the two groups. It therefore provides a strong control for any individual differences in memory abilities. These difference scores, relative to delayed span, are shown in Figure 2, and were subjected to a 2 (task: complex span/ Brown-Peterson) x 2 (processing group: verbal/nonverbal) mixed ANOVA. This revealed a significant main effect of group, $F(1, 28) = 6.23$, $p = .019$, $MSE = 2.13$, _^{2} = .182, due to lower scores (poorer performance relative to the delayed span task) in the verbal processing group. However, the task x processing group interaction was not significant: $F(1, 28) = 0.97$, $p = .332$, $MSE = 1.22$, _^{2} = .034.

Discussion

The aims of this initial experiment were to, first, compare the effects of verbal as opposed to non-verbal processing on working memory performance using directly comparable processing tasks, and, second, to examine whether the degree of forgetting caused by either type of processing varied across different types of working memory paradigm. A major strength of our design is that the participants in the two different processing groups carried out different types of processing on identical materials. Indeed, every participant in the experiment was presented with exactly the same computerised tasks, and these were not modified in any way at all to account for processing group membership. Rather, participants in the two groups were simply given different instructions as to how to process these common stimuli. In addition, both processing tasks required the same form of two-choice response, and we were partially successful in equating the two processing tasks for level of difficulty. Specifically, participants in the verbal processing group completed the baseline processing that took place both before and after the memory tasks slightly more rapidly than participants in the nonverbal processing group. Although this difference was not statistically significant, the cognitive load associated with making a correct response to processing was certainly no larger in the verbal than in the nonverbal processing group. Thus, it was not the case that verbally-mediated parity judgments required more time, and by inference attentional resource, to execute than did visual target identification.

Despite this, verbal processing led to greater forgetting than did nonverbal processing in this Experiment. Although the main effect of processing group on recall scores from the three memory tasks was not significant, it must be remembered that the delayed span task contained no processing, and so was identical for the two groups. As a result one would not expect a group difference on this task, and indeed this task serves as an excellent control for any potential individual differences in basic storage capacity. While the processing group by task interaction was close to significant when all three memory tasks were considered (p = .06), *an a priori* analysis of the two tasks in which processing was required, which controlled for individual differences in delayed span performance, clearly revealed a processing group effect. As Figure 2 shows, participants in the verbal processing group did less well on both the complex span and Brown-Peterson tasks than did individuals in the nonverbal processing group.

Figure 2 also demonstrates that participants within a given group did not perform at the same level on all tasks. At first sight the most striking aspect of Figure 2 might well be the finding that both groups of participants tended to score more highly on the complex span task than on the delayed span task. While this appears to suggest that the addition of processing had no effect on memory, it should be remembered that the structure of these two tasks is somewhat different. Given this, the most informative task comparisons are i) between delayed span and Brown-Peterson performance, where the structure of the tasks is identical except for the inclusion of processing in the latter, and ii) between Brown-Peterson and complex span tasks, where exactly the same total storage and processing demands are imposed on participants but the order of presentation of these demands differs.

Turning to the first of these comparisons, only individuals in the verbal processing group showed significantly poorer performance on the Brown-Peterson than on the delayed span task. This indicates that the type of nonverbal processing employed in this experiment was not demanding enough to cause any noticeable forgetting, despite the fact that, on average, nonverbal processing judgements took slightly longer to complete than verbal processing decisions. In contrast, both groups showed a significant difference in performance between the complex span and Brown-Peterson tasks, and this effect was comparable in the two samples. This clearly indicates that something about the interleaved structure of a complex span task makes it easier to perform than an analogous task in which a block of processing takes place after presentation of all of the storage items. We return to explanations for this effect in the General Discussion.

However, first we highlight one concern about the current data that potentially affects its interpretation. This is the fact that while there was no significant difference in the processing times between the two judgment types, accuracy for location judgments was significantly higher than parity judgments. It is worth noting that both processing groups achieved high levels of accuracy, and that even among individuals in the verbal processing group only 5% of responses were errors (compared to 2% in the nonverbal processing group). Nonetheless, it could be argued that a greater error rate for parity judgments interfered with attempts to maintain the memory items among individuals in the verbal processing group. A longstanding observation is that processing rate slows down after an error has been committed, possibly because individuals need time to evaluate and make adjustments to their responding (see Jentzsch & Dudschig, 2008). Recent research has shown that the decrement in recall, which was originally attributed to decay mechanisms (e.g., Portrat, Barrouillet, & Camos, 2008), may instead be caused by this type of "post-error" monitoring (Lewandowsky & Oberauer, 2009). Specifically, Lewandowsky and Oberauer argued that when difficult processing induces more errors, attention is deployed to the readjustment of responding, rather than to memory maintenance. This might therefore imply that the evidence of greater forgetting caused by verbal than nonverbal processing in this experiment follows from a higher rate of errors committed by the verbal than the nonverbal processing group.

An immediate counter to this concern is the fact that there was no evidence that processing rates were slower in the verbal processing group, contrary to what would be expected if errors led to slowing of subsequent processing judgements. Nevertheless, in the light of this issue Experiment 2 was conducted using a new set of processing stimuli, with the aim of equating both response time and accuracy for verbal and nonverbal processing judgements. Following a period of piloting, the stimuli used in Experiment 2 were letter pairs. Participants were again allocated to either the verbal (same-domain as storage) or nonverbal (different-domain to storage) processing group. In the former, participants had to decide if each letter pair rhymed, whereas those in the latter group had to judge if the two letters shared the same kind of symmetry (whether they were both vertically, or horizontally symmetrical). Apart from the materials adopted for the processing activity, the overall design and memory span tasks remained unchanged. As in Experiment 1, the time required to execute these types of processing was assessed using pre- and post- baseline tasks. Furthermore, the accuracy of the two types of processing was monitored, in order to ensure that the two types of processing produced comparable (and acceptable) levels of error rates. In this way, if recall is found to be impaired in the presence of same-domain processing, relative to different-domain processing, then this can confidently be attributed to domain-specific factors, rather to general differences in the level of difficulty of the two types of processing activity.

Experiment 2

Method

Participants

The participants were recruited at the University of Bristol, using the same age and native language criteria as in Experiment 1. None had participated in the previous experiment. They were either credited or reimbursed (£7) for their time. Data from five participants were discarded due to unacceptably high error rates (over 25%) in their processing. In total, data from 30 participants (3 males) were analyzed, all of whom were aged between 17 and 35 years, with 15 participants in each of the verbal and nonverbal processing groups.

Materials

The same storage items used in Experiment 1 were used here. That is, participants were presented with to-be-remembered objects both auditorily and pictorially. Also, as in

Experiment 1, each of the three memory span tasks (complex span, Brown-Peterson, and delayed span) consisted of trials containing 4, 5, 6, and 7 items, with four trials at each span length. However, in this study all participants were presented with letter pairs for the processing component of the tasks. Half of the letter pairs were rhyming (e.g., A and K, T and E), and half were non-rhyming (e.g., E and K, T and Y). Furthermore, of these letter pairs, half also contained members that shared an axis of symmetry (e.g., A and T both have an vertical axis, K and E both have a horizontal axis), whereas half did not (e.g., A and E, T and K). Each letter pair was presented with one letter on the left and the other on the right side of the screen in 84 point Arial font. For the baseline processing tasks, participants were presented with 4 letter pairs for practice, followed by 30 letter pairs from which data were recorded. There were an equal number of rhyming and non-rhyming, and symmetrical and non-symmetrical pairs in these stimuli sets.

Procedure

The overall procedure in Experiment 2 was identical to that in Experiment 1, with the exception that here, the processing component within the baseline and memory span tasks consisted of rhyme or symmetry judgments on letter pairs. Depending on their group membership, participants were told to press the “yes” key (“p” on the keyboard”) for letter pairs that rhymed (verbal processing) or shared a common axis of symmetry (nonverbal processing), and the “no” key (“q” on the keyboard) for letter pairs that did not rhyme or share a common axis of symmetry. As in Experiment 1, participants here were presented with a “before” and an “after” baseline task, and the three memory span tasks, with the order of presentation of these tasks counterbalanced across participants using a Latin square design. The experiment lasted approximately one hour in its entirety.

Results

Table 2 presents summary statistics for the performance of the two processing groups on the baseline processing and recall tasks.

Baseline Processing Tasks

RT. A 2 (task: pre-test/ post-test) x 2 (processing group: verbal/nonverbal) mixed ANOVA performed on the median correct RTs showed a significant main effect of task, $F(1, 28) = 105.58$, $p < .001$, $MSE = 49849$, _^{2} = .790, and a significant task x processing group interaction, $F(1, 28) = 20.52$, $p < .001$, $MSE = 49849$, _^{2} = .423. Overall, responding was faster in the post-test than the pre-test task, but the decrease in RT was more marked in the nonverbal than the verbal processing group. It is important to note though that there was no significant difference in RTs between the two processing groups in the post-test baseline measure, $t(28) = 0.89$, $p = .38$. The processing group main effect was also significant, $F(1, 28) = 5.16$, $p < .05$, $MSE = 118757$, _^{2} = .156. This effect arose because, averaged across sessions, more time was needed to make symmetry judgments than rhyme judgments.

Accuracy. The same mixed ANOVA was carried out to examine the accuracy levels achieved in the baseline tasks. Here, none of the main effects and interaction were significant. Of particular note is that the symmetry and rhyme judgments did not differ from each other in terms of the proportion of correct responses produced, $F(1, 28) < 1$, $p = .86$.

Memory Span Tasks

Recall scores. The method used to calculate a recall score for each participant on the three memory span tasks was the same as that employed in Experiment 1. The recall scores were then analyzed with a 3 (task: complex span/ Brown-Peterson/ delayed span) x 2 (processing group: verbal/nonverbal) mixed ANOVA. The main effect of task was significant, $F(2, 56) = 24.24$, $p < .001$, $MSE = 1.56$, _^{2} = .464, reflecting the fact that recall was on average lower in the Brown-Peterson task than in the complex span task, $F(1, 29) = 28.25$, $p = .001$, $MSE = 186$, _^{2} = .344, which in turn produced lower scores than the delayed span task, $F(1, 29) = 11.00$, $p = .017$, $MSE = 1.70$, _^{2} = .182. The main effect of processing group was not significant, $F(1, 28) < 1$, $p = .914$, but this was qualified by a significant task x processing group interaction, $F(2, 56) = 4.33$, $p = .018$, $MSE = 1.56$, _^{2} = .134.

As in Experiment 1, this interaction was explored by first exploring the effect of task in each processing group separately. Individuals in the verbal processing group showed poorer performance on the Brown-Peterson task than on either the delayed span task, $t(14) = 6.21$, $p < .001$, or the complex span task, $t(14) = 5.23$, $p < .001$. Individuals in the nonverbal processing group also had lower recall scores on the Brown-Peterson than on the delayed span task, $t(14) = 4.33$, $p = .001$, but there was no significant difference in this group's performance on the Brown-Peterson and complex span tasks, $t(14) = 1.22$, $p = .241$. A subsequent analysis of difference scores relative to delayed span was performed (see Figure 3). This revealed a main effect of group that was close to significant, $F(1, 28) = 3.75$, $p = .063$, $MSE = 4.46$, _^{2} = .118, again reflecting greater forgetting relative to delayed span in the verbal processing group. In addition, the task x processing group interaction was significant, $F(1, 28) = 4.85$, $p = .036$, $MSE = 1.64$, _^{2} = .148. As Figure 3 suggests, the two processing groups did not differ significantly in the degree of forgetting, relative to delayed span, on the complex span task, $t(28) = 0.48$, $p =$

.636. However, the verbal processing group showed a significantly larger relative impairment than the nonverbal group on the Brown-Peterson task, $t(28) = 3.02$, $p = .006$.

Discussion

The first point to make about this second Experiment is that the adoption of a new set of stimuli for the processing component of the Brown-Peterson and complex span tasks led to more similar levels of accuracy for the two types of processing than was seen in Experiment 1. In this case there was no suggestion that verbal processing (rhyme judgements) led to more errors than nonverbal processing (symmetry judgements), indeed average error rates were extremely similar and very low (2.8% for verbal processing vs. 2.6% for nonverbal processing). One might suggest that such low error rates indicate that participants were at ceiling on these processing tasks, and that an underlying difference in difficulty between the two types of processing would be seen if the tasks were made harder. Certainly these processing judgements took somewhat less time than the corresponding judgements in Experiment 1 (compare Tables 1 and 2). However, the RT data from the baseline processing tasks clearly indicate that verbal processing was no harder than nonverbal processing in this experiment. To the contrary, on average individuals in the verbal processing group completed processing more rapidly than their counterparts in the nonverbal processing group, although this group difference was only significant at the pre-test and not the post-test.

Despite this, and the fact that once again participants in the two groups were shown entirely identical processing stimuli, the type of processing carried out on these stimuli did have an effect on performance in the Brown-Peterson and complex span tasks. As shown in Figure 3, verbal processing led to significantly greater forgetting in the

Brown-Peterson task than did nonverbal processing. In contrast, and in a departure from the pattern seen in Experiment 1, the two types of processing led to comparable performance on the complex span tasks. It is worth noting that while the degree of forgetting shown by the nonverbal processing group on the Brown-Peterson task, relative to the delayed span task, was less than that seen in the verbal processing group, these individuals still showed significantly poorer performance on the Brown-Peterson task than on the delayed span task. This is in contrast to Experiment 1, and suggests that nonverbal processing can have a detrimental effect on working memory, in line with the findings of Barrouillet, Vergauwe, and colleagues, although clearly this effect is less marked than the effect of verbal processing in this instance. Among individuals in the verbal processing group Brown-Peterson performance was lower than that seen on complex span, while in the nonverbal group recall scores for these two tasks were comparable.

The results of this second experiment therefore mirror those of Experiment 1 in some regards, but differ in others. Before discussing the overall pattern of data we present details of a third Experiment that was carried out to address a potential concern with the first two experiments, namely that storage items in all tasks were presented both auditorily and pictorially. This dual mode of presentation was adopted because we wanted to ensure that participants accurately encoded each storage item, particularly given the use of an open set of memoranda. Pictures were used to disambiguate any potentially hard to identify items because these tasks had initially been developed for use with children. However, it is possible that participants in the first two experiments did not encode and maintain storage items in a phonological code to the degree that they might have, and instead may have maintained these items on the basis of their visual representations. If so, then this would attenuate any potential domain-specific effects. Experiment 3 addressed this issue by replicating Experiment 2, with the exception that the storage items were presented only in their auditory forms.

Experiment 3

Method

Participants

The participants were recruited from the University of Bristol in the same manner as in Experiments 1 and 2. None had participated in previous experiments and were credited or reimbursed (£7) for their participation. After excluding data from four individuals (due to the high levels of error rates in their processing judgments), data from 30 participants (13 males) were analyzed. There were 15 participants in each of the verbal and nonverbal processing groups. Once again, all participants fell within the 17 to 35 year old age range.

Materials

These were identical to those used in Experiment 2.

Procedure

The procedure was as described in Experiment 2, with the exception that across all three memory span tasks (complex span, Brown-Peterson, delayed span), the storage items were presented only in their auditory forms, rather than both pictorially and auditorily in the previous two experiments.

Results

Summary statistics for the two processing groups' performance on the baseline processing tasks and the memory tasks is given in Table 3.

Baseline Processing Tasks

RT. The median RT obtained from each participant, for the pre- and post-test baseline processing tasks, were analyzed in the same way as in previous experiments. Participant were significantly faster at responding pre-test than post-test, $F(1, 28) = 89.20$, $p < .001$, $MSE = 51055$, $\_^2 = .761$. The pre- and post-test difference in RT was greater for the nonverbal than the verbal processing group, $F(1, 28) = 23.80$, $p < .001$, $MSE = 51055$, $\_^2 = .459$. However, a non-significant main effect of processing group, , $F(1, 28) = 1.02$, $p = .322$, showed that there was no overall difference in the RTs between the two processing groups. Indeed, average RTs were numerically smaller in the verbal processing group.

Accuracy. Both groups produced high proportions of correct responses. The analysis revealed no significant main effect of processing group, $F(1, 28) = 3.56$, $p = .070$, although there was a trend for numerically superior processing accuracy in the nonverbal processing group. In addition, neither the task main effect nor the task x processing group interaction was significant, $F(1, 28) = 2.85$, $p = .10$ and $F(1, 28) < 1$, $p = .90$ respectively.

Memory Span Tasks

Recall scores. Recall scores were calculated and then analyzed in the same manner as in previous experiments. The task main effect was significant, $F(2, 56) = 38.05$, $p < .001$, $MSE = 1.49$, $\_^2 = .576$. Overall, recall was significantly worse in the

Brown-Peterson than in the complex span task, $F(1, 29) = 41.67$, $p < .001$, $MSE = 1.36$, _^{2} = .590, which in turn produced significantly poorer performance than the delayed span task, $F(1, 29) = 4.28$, $p = .048$, $MSE = 1.74$, _^{2} = .129. The main effect of processing group was not significant, $F(1, 28) = 0.24$, $p = .629$, but this was again qualified by a significant task x processing group interaction, $F(2, 56) = 3.93$, $p = .025$, $MSE = 1.49$, _^{2} = .123.

This interaction was again explored by first examining task effects in each processing group. Members of the verbal processing group showed poorer performance on the Brown-Peterson task than on both the delayed span task, $t(14) = 7.71$, $p < .001$, and the complex span task, $t(14) = 4.39$, $p = .001$. Individuals in the nonverbal processing group similarly performed less well on the Brown-Peterson task than on both the delayed span, $t(14) = 4.14$, $p = .001$, and complex span tasks, $t(14) = 4.84$, $p < .001$. A further analysis of difference scores relative to delayed span was again performed (see Figure 4). This revealed a significant main effect of group, $F(1, 28) = 6.93$, $p = .014$, $MSE = 4.75$, _^{2} = .198, due to greater forgetting relative to delayed span in the verbal processing group. However, the task x processing group interaction was not significant, $F(1, 28) = 0.50$, $p = .486$, $MSE = 1.39$, _^{2} = .018.

Discussion

Experiment 3 differed from Experiment 2 solely in terms of the method of presentation of the to-be-remembered storage items; in this case presentation of the memoranda was auditory only. A comparison of Tables 2 and 3 shows that this change had a small but noticeable effect on overall levels of performance on the three memory tasks, leading to a drop of just over 1 point on the recall score scale. This is an indication that participants were engaging in some form of dual coding of the memoranda in Experiment 2 (and by implication Experiment 1 as well), although perhaps not to any great extent. However, a corresponding comparison of Figures 3 and 4 indicates that the degree of forgetting caused by the imposition of verbal processing in the complex span and Brown Peterson tasks was broadly comparable in Experiments 2 and 3. In other words, removing visual support at the point of presentation of the storage items did not lead to verbal processing being much more detrimental to recall.

Having said this, the overall pattern of results from Experiment 3 is not identical to that seen in Experiment 2, and in fact is closer in form to that seen in Experiment 1. A significant interaction between processing group and task was again observed when performance was considered across all three tasks, but when the two groups were compared on the two tasks that involved processing (complex span and Brown-Peterson), controlling for any differences in delayed span performance (see Figure 4), individuals in the verbal processing group showed a significantly greater effect of the imposition of processing than did individuals in the nonverbal processing group on both of these tasks. In addition, participants in both processing groups performed less well on the Brown-Peterson task than on either the delayed span or complex span tasks.

One final point to note about Experiment 3 is that was a trend (p = .07) for more errors to be made during baseline processing in the verbal than the nonverbal processing group. Clearly there was no reason to expect this trend in advance, given that the processing employed in this Experiment was identical to that used in Experiment 2 where the two groups showed comparable levels of accuracy. In addition, as in all experiments levels of baseline processing accuracy were very high in both groups (on average 95% in the verbal group and 97% in the nonverbal group), and so relatively few errors were made even in the verbal processing group. We return to this point in the General Discussion, but note here that, as in Experiment 1, there was no evidence at all that any potential difference in error rates led to a disadvantage in terms of response times in the verbal processing group. Post-error attentional processing would necessarily slow subsequent responses, but RTs for the verbal processing group were numerically (though non-significantly) lower than those for the nonverbal processing group.

Analysis of additional data drawn from Experiments 2 and 3

The fact that Experiments 2 and 3 differed solely in terms of the presentation mode of the storage items allowed for a comparison of additional aspects of performance from these studies that benefited from the increased power gained from collapsing across them. In particular we first explored the question of whether participants slowed their processing relative to baseline levels as a result of the memory demands involved in the complex span and Brown-Peterson tasks, in order to look for any evidence of the strategic delaying of processing that would accompany either refreshment or rehearsal activities. The analyses of RTs in pre- and post-test baseline sessions in each experiment showed that participants' response times decreased over the course of each experiment, particularly so for nonverbal processing in Experiments 2 and 3. This is unsurprising and no doubt reflects practice effects. However, to take this into account we fitted the RT learning curve shown by each participant across the course of Experiments 2 and 3. From this we examined whether RTs for one of the processing-loaded memory tasks differed from that expected given both baseline performance and the RTs seen in the other processing-loaded memory task.

The decision of exactly how to perform this analysis was affected by the counterbalancing employed in the above experiments. As already noted, a Latin square design was used to create three different orders of presentation of the three memory tasks. The particular version of counterbalancing employed meant that two thirds of each sample received the Brown-Peterson task before the complex span task while only one third received these two tasks in the reverse order. Given this, we elected to use processing reaction times on the Brown-Peterson task, along with the two measures of baseline processing RT, to fit the learning curve for each participant, based on the assumption that the greatest variation in RT would be in the earlier parts of the overall testing session. As each memory task employed a span procedure, and successively presented trials at storage lengths 4, 5, 6 and 7 in that temporal order, four data points for RTs were taken from each of the Brown-Peterson and complex span tasks by averaging the median processing RTs derived from trials at a given span length. For the Brown-Peterson task the median RT for all correct processing responses completed in the single block of processing was taken from each trial. For the complex span task the median RT for all correct processing responses within each segment of processing was taken, and then these were averaged across the processing segments (4 for span length 4, 5 for span length 5 etc., see Figure 1) for each trial.

A learning curve for each participant was therefore fitted to the pre-test baseline processing RT, the four successive RT measures from each level of the Brown-Peterson task, and the post-test baseline processing RT, using each of these phases of the overall experiment as a surrogate for time (see Figure 5). Initial analysis showed that power functions provided the best fit to these data (cf. Newell & Rosenbloom, 1981), and the resultant parameters for each individual's power function were used to derive expected values for their four complex span RTs, which in turn were compared to the observed complex span RTs to derive residual scores for these data points. Any significant deviation from the expected values for the complex span task would indicate that participants either speeded or slowed their processing responses relative to the Brown-Peterson task and pre- and post-test baseline sessions.

Figure 5 plots the actual RT data obtained from each section of each task for the two processing groups, plus the average power curve function fitted through the baseline and Brown-Peterson data points. Individuals' residual scores for the four levels of the complex span task, based on each own individual's power curve function, were subjected to a one-sample t-test. In the verbal processing group the residual RTs for complex span levels 4 and 5 were not significantly different from zero, $t(19) = 0.891$, $p = .383$, $t\,(19) = 0.479$, $p = .637$ respectively. However, the residual RTs for levels 6 and 7 were significantly greater than zero, $t(19) = 4.080$, $p < .001$, $t\,(19) = 3.185$, $p = .004$ respectively. Similarly, in the nonverbal processing group residual scores for complex span did not differ significantly from zero at levels 4 and 5, $t(19) = 0.170$, $p = .866$, $t\,(19) = 1.631$, $p = .119$, but were significantly greater than zero at levels 6 and 7, $t(19) = 3.171$, $p = .005$, $t\,(19) = 3.885$, $p < .001$. A final analysis compared the size of the residual scores associated with each processing group at each level of the complex span task, and showed no evidence of any significant group differences in any of these residuals between participants carrying out verbal or nonverbal processing, $p \geq .399$.

These findings suggest that individuals selectively slowed their processing during the higher span levels of the complex span task, relative to their Brown-Peterson performance. In addition, this degree of slowing was comparable in the two different processing groups. This is not to say that individuals did not also slow their processing in the Brown-Peterson task, because as Figure 5 suggests the average best-fit power curves do slightly overestimate both baseline processing RT values in each group. However, what is clear is that any slowing that takes place is greater in the complex span than in the Brown-Peterson task.

To look at this issue in more detail a final set of analyses examined the degree of any slowing that occurred across the various ‘processing positions’ of the complex span trials (cf. Friedman & Miyake, 2004; Saito & Miyake, 2004; Towse, Hitch, & Hutton, 2000). As an episode of processing followed the presentation of each stimulus in the complex span task, the number of processing episode in each trial varied in line with list length; in this case from 4 to 7. Given that the processing tasks employed in Experiments 2 and 3 were identical, these analyses were again carried out on the combined sample of 60 individuals who took part in these two studies. Two measures were taken from each type of processing episode across list lengths and the position of the processing episode in that list: i) the average reaction time for the individual’s first response in that interval, and ii) the average of all subsequent reaction times in that interval. This allowed us to examine whether any slowing of responses was restricted to the first processing response, which would be indicative of rehearsal, refreshment or consolidation activities being interpolated by the participant before starting to make processing responses.

A series of preliminary 2 (experiment: 2/3) x 2 (processing group: verbal/nonverbal) x list length (4 to 7) mixed ANOVAs were performed on both first and subsequent processing responses at each list length. These showed that experiment (2 or

3) did not interact significantly with processing episode position for analyses at list lengths 5, 6 or 7, $F \geq 2.08$, $p \leq .085$ (significant interactions between experiment and processing episode position were observed for first responses but not for subsequent responses at list length 4). The main effect of processing group, and the processing group by episode position interaction, from each of these analyses are summarised in Table 4. The table shows that while processing group interacted significantly with position for both dependent measures at list length 4, there were no significant interactions at longer list lengths. Consequently, for the sake of clarity the analyses reported below collapsed data across experiment and processing group to look at the effects of processing episode position on both first and subsequent processing response times in list lengths 5 to 7.

Figure 6 plots the average correct reaction times for both first responses and subsequent responses within each processing position of each of these complex span list lengths. What is immediately apparent from the figure is that first responses were substantially longer than subsequent responses; further analyses therefore continued to treat these two dependent measures separately. First, the main effect of processing position for each of these measures was decomposed into its linear and quadratic trends, with separate analyses examining these trends at each list length. The results of these analyses are summarised in Table 5, which shows that strong linear trends were apparent for each measure at each list length, with additional reliable quadratic trends for first responses at list length 7 and for subsequent responses at list lengths 6 and 7.

These results, and the data shown in Figure 6, are suggestive of a linear increase in response times across the first four or five episode positions, followed by a levelling off of these times between positions 5 and 7. Indeed, an analysis that contrasted change in reaction times across positions 2 and 4 with that seen across positions 5 and 7 of list length 7 showed a significant increase in first response times between the former pair, $F(1, 58) = 45.03$, $p < .001$, $MSE = 115676$, _^{2} = .437, but not between the latter pair of positions, $F(1, 58) = 2.17$, $p = .146$, interaction: $F(1, 58) = 29.54$, $p < .001$, $MSE =$ 138790, _^{2} = .337. A similar pattern was seen for the corresponding analysis of subsequent response times, with a significant increase between positions 2 and 4, $F(1, 58) = 4.68$, $p = .035$, $MSE = 19826$, _^{2} = .073, but not between positions 5 and 7, $F(1, 58) = 0.05$, $p = .824$; however, in this case the interaction between these effects was not significant, $F(1, 58) = 2.75$, $p = .103$.

These data therefore show that the evidence of relative slowing of processing times in the longer complex span list lengths that emerged from the learning curve analyses depicted in Figure 5 arises for two related reasons. First, all reaction times within a given processing episode tend to increase with episode position (cf. Friedman & Miyake, 2004; Saito & Miyake, 2004). This effect is not purely linear, and there is clearly evidence of a levelling off of this increase at the latter episode positions of longer lists, which we return to below. However, the fact that longer lists therefore contain relatively more episodes with ‘long’ responses accounts for the degree of overall slowing seen on these lists. Second, while this effect is seen in the subsequent responses that follow the initial response in each episode position, it is much more marked in these first responses. Indeed, the average increase in response time for first responses across positions 1 to 5 at list lengths 5 to 7 was 306 ms (95% CI = 208 to 404 ms). This contrasts to an average increase in subsequent responses across these positions on these lists of 101 ms (95% CI = 71 to 131 ms). Consequently, the major reason for the overall slowing of processing times on the longer complex span tasks is that participants delay their *first* response to processing more at later positions in a complex span list.

General Discussion

The main aim of the three experiments presented above was to clarify the question of whether the effects of processing on recall in working memory paradigms are domain-general or domain-specific. A secondary aim was to determine whether the degree of domain-generality or specificity observed depended on the nature of the working memory task employed. A key strength of all three experiments, which represents an important advance on all previous work, is that participants in the two processing groups in each study carried out either ‘verbal’ or ‘nonverbal’ processing on identical materials. In addition, we made every attempt to ensure that the verbal processing involved in each study, which one might expect to lead to greater forgetting of the kind of verbal memoranda employed here, was no more difficult than the non-verbal processing.

It should, of course, be noted that individuals in the verbal processing group in Experiment 1 produced more errors during the baseline processing tasks than did their counterparts in the nonverbal processing group. This prompted a change in the processing materials for Experiments 2 and 3, but a similar, though non-significant, trend was observed in Experiment 3. However, in all experiments error rates were very low. Furthermore, while one might argue that relatively increased error rates for the verbal processing groups in Experiments 1 and 3 might lead to an increase in the cognitive load associated with verbal as opposed to nonverbal processing, there was no evidence in any experiment that verbal processing took longer to complete than nonverbal processing (see also Figure 5 for processing times within the memory tasks and Table 5 for a direct comparison of processing times by processing type in the complex span task), which is what one would predict if increased errors in the verbal processing groups led to noticeable post-error recruitment of attention and slowing of subsequent responses (Lewandowsky & Oberauer, 2009). Consequently, we would argue that we were successful in ensuring that the cognitive load associated with verbal processing was no more demanding than that associated with nonverbal processing in each experiment.

Despite this, the clearest finding from the three studies is that verbal processing led to greater forgetting than did nonverbal processing. In all three experiments individuals in the verbal processing group did significantly less well on the Brown-Peterson task than those in the nonverbal processing group, once baseline memory performance on the delayed span task had been taken into account (see Figures 2 to 4). Given that there are good reasons to argue that verbal processing was no more attentionaly demanding than nonverbal processing this clearly indicates that something about the nature, rather than the difficulty, of the verbal processing led to increased forgetting. In addition, in Experiments 1 and 3 verbal processing led to poorer performance on the complex span task than did nonverbal processing.

However, it is important to note that nonverbal processing does also lead to forgetting in these experiments. The design of our studies was such that one can make a particularly informative comparison between the delayed span and Brown-Peterson tasks, as these tasks have exactly the same structure but differ in terms of whether a processing load is imposed or not. In Experiments 2 and 3 participants in the nonverbal processing group performed significantly less well on the Brown-Peterson than on the delayed span task. While the degree of forgetting caused by this processing load was less than that seen among the verbal processing group, this clearly shows that engaging in nonverbal judgements that require a two-choice response decision can affect recall of verbal information. Having said this, in Experiment 1 nonverbal processing did not lead to reliable forgetting in the Brown-Peterson relative to the delayed span task. One possible explanation for this is that the nonverbal processing employed in this first experiment was based on single target detection, rather than a judgement about the nature of two related stimuli (as in Experiments 2 and 3). The target detection involved in Experiment 1 was certainly time consuming, as average nonverbal processing times for this Experiment were at least as long as for Experiments 2 and 3 (compare Tables 1 to 3). However, the nonverbal processing in Experiment 1 essentially involved a serial, if difficult, visual search for a target among homogeneous distractors, followed by a decision about the target's position in space once it was detected. It is therefore possible that while the search component of this processing is time consuming, it is not particularly attentionally demanding in the way that it would be if heterogeneous distractors were present to capture attention (e.g., Lavie & de Fockert, 2006). If so, then the majority of the RT associated with these processing judgements would reflect non-attentional search processes and only the minority of the time would be taken up by an attentionally-demanding forced choice decision.

The other particularly meaningful comparison between tasks that can be made in each experiment is between the Brown-Peterson and complex span tasks, which share the same total storage and processing demands but which present these demands in different temporal formats. Generally speaking performance was better in the complex span tasks than the Brown-Peterson tasks employed here. In Experiment 1 this difference was only close to significant for the verbal processing group (*p* = .06) while being clearly significant in the nonverbal group. In Experiment 2 this effect was only reliable in the verbal processing group, while in Experiment 3 this difference was significant in both processing groups. There seems no obvious explanation for this slight discrepancy in results, and our reading of the findings is that they indicate a general tendency for complex span performance to be superior to Brown-Peterson performance in both groups. Indeed, this finding fits with data from two experiments conducted by Tehan et al. (2001) in which adults' performance on complex span and Brown-Peterson tasks at span length 4 were compared.[2] In that study participants tended to show better complex span than Brown-Peterson performance, even though the storage and processing demands of the two tasks were well matched, as here. However, while this effect emerged in a number of the analyses conducted by Tehan et al., it was not significant in every comparison. A slightly different pattern of findings emerged from an earlier study in our lab employing similar tasks with children (Tam, Jarrold, Baddeley, & Sabatos-De-Vito, 2010) in which 6- and 8-year-old children showed no significant differences between levels of performance on directly comparable Brown-Peterson and complex span tasks.

The data from our additional analyses of processing times within the working memory tasks (see Figures 5 and 6) may well be relevant to the question of when complex span tasks produce significantly higher levels of performance than comparable Brown-Peterson tasks. These data and the corresponding analyses showed that participants in both processing groups slowed their processing responses in complex span tasks, relative to Brown-Peterson tasks, although only measureably so at the higher list lengths (span lengths 6 and 7). This in turn reflected a general trend for processing responses to increase in latency with episode position in the complex span trial; in other words, responses made in the later processing episodes after the majority of memoranda had been presented were slower than those from earlier processing episodes. This effect was observed for all processing responses, but was particularly marked for the first response within each episode. This suggests that participants were delaying their responses to engage in some form of maintenance-related activity, and the fact that these responses increased in latency with episode position is consistent with the suggestion that at least a subset of the previously presented storage items were the subject of that activity, rather than just the immediately previous storage item. The fact that longer complex span trials necessarily contain more processing episodes, that were therefore subject to relatively greater slowing of responses, accounts for the evidence of overall slowing seen on longer complex span trials in Figure 5.

However, Figure 6 also indicates that the degree of slowing is not entirely linear across episode position, but rather levels-off at particularly late positions. If we are right in assuming that the linear increase in response times across early episode positions reflects maintenance strategies being applied to a number of memoranda, then this implies that there was a limit to the number of storage items that participants applied these strategies to. This might be because of some form of capacity limit (cf. Baddeley & Hitch, 1974; Cowan, 2001), or because participants are reluctant to delay their first processing response for too long. Regardless, this suggests that one reason why the same degree of slowing is not apparent in Brown-Peterson tasks is that in that paradigm participants are presented with all the memoranda in a single block at the start of the trial. On longer lists the number of to-be-remembered storage items may well be greater than the number of items they could apply maintenance strategies to, or the time this would take may be longer than they are prepared to devote to that task. Another, related reason why this form of slowing is less likely in Brown-Peterson tasks is that the blocking of processing may well 'entrain' the participant into focussing solely on these processing operations. In contrast, the fact that processing operations are split into discrete segments in the complex span tasks imposes switches between storage and processing on the participant, which may well make them more likely to engage in their own 'switches away' from processing to maintenance-related activity.

The current data therefore suggest that one form of maintenance activity is more likely in complex span than in Brown-Peterson tasks, although this task difference may be particularly apparent when relatively long list lengths are employed in these procedures. In turn this provides a potential explanation for both the generally higher complex span than Brown-Peterson recall scores in this study and the fact that this difference was not always consistently observed (cf. Tehan et al., 2001); the above analysis suggests that a larger task effect may have been apparent if all trials on both tasks had employed relatively long list lengths. [3] In that regard it is worth noting that our earlier work that failed to find this task difference in children employed a span procedure, rather than the presentation of potentially supra-span lists (Tam et al., 2010).

This raises the question of exactly what these maintenance activities involve. The present data do not answer this question directly, but the pattern of findings suggest that individuals engage in attentional refreshment in the intervals between processing responses, and particularly so at the start of the processing episode. The fact that initial processing response time increased linearly across initial episode positions is not consistent with the suggestion that participants are consolidating only the just-presented item into long term memory; rather, this suggests that participants are applying some form of maintenance activity to an ever increasing set of items. However, the extent of observed slowing, in the order of 60 ms per item assuming all items are subjected to this activity, is almost certainly too small to reflect subvocal rehearsal of the memoranda. Of course, not all participants will necessarily be engaging in strategic slowing on every processing episode, but previous estimates of the rate of subvocal rehearsal suggest that monosyllabic words of the form employed here are likely to take around 400 to 500 ms each to rehearse (Baddeley, Thomson, & Buchanan, 1975; Standing, Bond, Smith, & Isely, 1980; Thorn & Gathercole, 2001). In addition, one would expect verbal processing to be more likely to block rehearsal than nonverbal processing, to the extent that rehearsal might be possible during nonverbal processing. If so, then only participants in the verbal processing group would need to selectively slow processing to engage in rehearsal. This was not was observed, rather the degree of slowing of processing in complex span tasks was comparable in the two processing groups.

In contrast, refreshment is assumed to operate more rapidly than subvocal rehearsal as it need not involve the retrieval of the full phonological form of the memoranda (Barrouillet et al., 2007; Cowan et al., 1998; Hudjetz & Oberauer, 2007; Raye, Johnson, Mitchell, Reeder, & Greene, 2002). A number of authors have suggested that refreshment operates only on the most recently presented item (Raye et al., 2002; Oberauer & Lewandowsky, 2008). However, Cowan et al. (1998) argued for a form of reactivation in working memory that shares similarities with the ‘memory scanning’ processes that potentially operate in the Sternberg scanning task (Sternberg, 1966), where participants have to determine whether a target item was present in a stimulus set of a varying length of items. Sternberg's (1966) original data from this paradigm suggested that participants scan through the stimulus set at a rate of between 25 and 30 ms per item. This raises the possibility that refreshment might operate on more than one item, and the values obtained from these scanning tasks are certainly broadly in line with the estimate of refreshment rate obtained here. In addition, refreshment is assumed to be blocked by any processing that requires attentional retrieval from long-term memory, and hence would be expected to be compromised equally by the two types of processing employed in our studies. We therefore suggest that our data indicate that refreshing of multiple memoranda is taking place in our complex span tasks, to an equal extent in both processing groups, but to a greater degree than seen in the Brown-Peterson paradigms.

The current experiments therefore provide evidence of a domain-general effect of processing on recall within working memory paradigms, which we attribute to the disruption of refreshment. However there are clearly domain-specific effects of processing operating in addition. As noted above, although nonverbal processing did tend to produce a recall decrement, the effect of verbal processing on performance was more marked, and particularly so on the Brown-Peterson task employed in each experiment. At the outset of the paper we outlined two possible explanations for this effect, one being that verbal processing selectively blocks rehearsal, the other that verbal processing representations interfere more with the maintenance of verbal memoranda than do non-verbal processing representations. The current data do not unequivocally decide between these two suggestions (and indeed, they are notoriously difficult to distinguish, Lewandowsky et al., 2009), but certain versions of an interference account seem not to be consistent with the current data. In particular, the fact that the verbal processing tasks either generated parity judgements about numbers (Experiment 1) or rhyme judgements about pairs of letters (Experiments 2 and 3) makes it unlikely that the representations generated by these tasks would provide substantial response interference at the point of recall (cf. Kane & Engle, 2000). As the memoranda employed were concrete words it seems unlikely that the generated verbal processing representations would be strong recall competitors (cf. Conlin & Gathercole, 2006; Oberauer, 2009). However, it remains possible that these generated verbal representations would have interfered with, or overwritten, the representations of memoranda, given that the two classes of items shared phonological features with each other (Farrell & Lewandowsky, 2002; Nairne, 1990; Oberauer, 2009; Oberauer & Lewandowsky, 2008). Certainly, this could readily explain why verbal processing leads to greater forgetting than nonverbal processing in the current tasks, even though verbal processing took no longer to complete.

At the same time, this finding is equally well explained by the blocking of rehearsal account, as verbal processing would undoubtedly make subvocal rehearsal more difficult, than would nonverbal processing.[4] Indeed, rehearsal may well be completely impossible during verbal processing of the form employed here, where each successive processing operation follows on directly from the completion of the previous one (Hudjetz & Oberauer, 2007). Given this, one might question why verbal processing does not lead to a greater degree of forgetting than currently observed. A possible explanation for this is that individuals are unlikely to be able to accurately rehearse all of the items presented on the longest length Brown-Peterson tasks. In recent, related, work we have shown that verbal processing of the form employed here leads to a greater decrement to recall in Brown-Peterson-like tasks than does nonverbal processing, but only up to a certain list length. In that study participants showed evidence of rehearsing up to four 2-syllable words during nonverbal processing, but no clear evidence of rehearsing lists of 5 to 7 items (Jarrold, Tam, Baddeley, & Harvey, 2010). This suggests that in the Brown-Peterson tasks employed here, participants in the nonverbal processing groups may have attempted rehearsal on shorter list length trials, but not on longer list length trials, given the danger that anything other than perfect rehearsal would dramatically impair recall (Lewandowsky et al., 2008; Oberauer & Lewandowsky, 2008). Averaged across all list lengths, this would reduce the advantage to performance, relative to that shown by verbal processing groups, offered by the opportunity to rehearse.[5]

Regardless of the precise cause of this apparently domain-specific effect of processing on storage performance, it is clear that the *nature* of the processing involved in a working memory task can affect performance, over and beyond any effects of the *difficulty* of that processing as measured in terms of its cognitive load. The current data therefore call into question the original assumptions of the TBRS model (e.g., Barrouillet et al., 2004; 2007), but are not inconsistent with the most recent instantiations of that account (e.g., Camos et al., 2009). Camos et al. (2009) suggest that processing can disrupt performance in working memory tasks by either blocking refreshment or rehearsal. The three experiments presented here provide novel support for this general class of model in that they demonstrate both domain-general and both domain-specific effects of processing, potentially due to disruption of refreshment and rehearsal respectively, using very carefully matched storage and processing tasks. In addition, the current study goes beyond previous studies in suggesting that the balance of these effects varies across working memory paradigm. Specifically, the importance of refreshment is most clearly seen on complex span tasks, where the imposed switches between processing and storage phases of the task prompt refreshment at the start of each processing episode, while the influence of rehearsal (or interference) is more apparent on Brown-Peterson tasks of the form employed here. This provides a potential explanation for why domain-general effects of processing appear to be seen more often seen on complex span tasks than on dual task paradigms (see Introduction). What remains to be addressed is whether this difference between paradigms has implications for their relative predictive power in terms of correlations with measures of intelligence and academic ability. Given that complex span tasks tend to be strong correlates of intelligence and academic attainment, the intriguing question that arises is whether Brown-Peterson-type tasks of the form employed here might be even more strongly related to these important 'real-world' measures (cf. Unsworth & Engle, 2006; 2007).

References

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

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

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

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.

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*, 570-585.

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology, 45*, 477-490.

Bayliss, D. M., Jarrold, C., Baddeley, A., & Gunn, D. M. (2005). The relationship between short-term memory and working memory: Complex span made simple? *Memory, 13*, 414-421.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General, 132,* 71-92.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539-576.

Brown, J. (1958). Some tests of the decay theories of immediate memory. *The Quarterly Journal of Experimental Psychology, 10*, 12-21.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205-228.

Bunting, M. F. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 183-196.

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

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33,* 386-404.

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory and Cognition, 30*, 1086-1095.

Colom, R., Rebello, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory and Cognition, 34*, 158-171.

Conlin, J. A., & Gathercole, S. E. (2006). Lexicality and interference in working memory in children and in adults. *Journal of Memory and Language, 55*, 363-380.

Conlin, J. A., Gathercole, S. E., & Adams, J. W. (2005). Stimulus similarity decrements in children's working memory span. *The Quarterly Journal of Experimental*

Psychology, 58A, 1434-1446.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163-183.

Conway, A. R. A., 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 and Review, 12,* 769-786.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87-114.

Cowan, N., Elliot, E. M., Saults, J. S., Morey, C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive abilities. *Cognitive Psychology, 51*, 42-100.

Cowan, N., Johnson, T. D., & Saults, J. S. (2005). Capacity limits in list item recognition: Evidence from proactive interference. *Memory, 13*, 293 - 299.

Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*, 141-160.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

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

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.

Duff, S. C., & Logie, R. H. (2001). Processing and storage in working memory span. *The Quarterly Journal of Experimental Psychology, 54A*, 31-48.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309-311.

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *The Quarterly Journal of Experimental Psychology, 38A*, 675-688.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin and Review, 9*, 59-79.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51*, 136-158.

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 556-569.

Hale, S., Myerson, J., Rhee, S. H., Weiss, C. S., & Abrams, R. A. (1996). Selective interference with the maintenance of location information in working memory. *Neuropsychology, 10*, 225-240.

Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory and Cognition, 35*, 1675-1684.

Hutton, U. M. Z., & Towse, J. N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory, 9*, 333-348.

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2010). The nature and the position of processing determines why forgetting occurs in working memory tasks *Manuscript submitted for publication.*

Jarrold, C., & Towse, J. T. N. (2006). Individual differences in working memory. *Neuroscience, 139*, 39-50.

Jentzsch, I. & Dudschig C. (2008). Why do we slow down after an error? Mechanisms underlying the effects of post-error slowing. *The Quarterly Journal of Experimental Psychology, 62*, 209-218.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 336-358.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 66-71.

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.

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General, 133*, 355-381.

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

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex memory spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 1118-1133.

Lavie, N., & de Fockert, J. (2006). Frontal control of attentional capture in visual search. *Visual Cognition, 14*, 863-876.

Lépine, R., Bernadin, S., & Barrouillet, P. (2005). Attention switching and working memory spans. *European Journal of Cognitive Psychology, 17*, 329-345.

Lewandowsky, S., Geiger, S. M., & Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *Journal of Memory and Language, 59*, 200-222.

Lewandowsky, S., & Oberauer, K. (2009). No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1545-1551.

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

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove: Lawrence Erlbaum Associates.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologia, 75*, 55-74.

Maehara, Y., & Saito, S. (2007). The relationship between processing and storage in working memory span: Not two sides of the same coin. *Journal of Memory and Language, 56*, 212-228.

Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 703-711.

Nairne, J. (1990). A feature model of immediate memory. *Memory and Cognition, 18*, 251-269.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Oberauer, K. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition, 37*, 346-357.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference. *Psychological Review, 115*, 544-576.

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.

Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning and Verbal Behavior, 10*, 346-354.

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

Portrat, S., Barrouillet, P., & Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1561-1564.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a single thought: Dorsolateral pfc activity associated with refreshing just-activated information. *NeuroImage, 15*, 447-453.

Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language, 59*, 1-17.

Saito, S., & Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in reading span performance. *Journal of Memory and Language, 50*, 425-443.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4-27.

Smyth, M. M., & Pelky, P. L. (1992). Short-term retention of spatial information. *British Journal of Psychology, 83*, 359-374.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174-215.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*, 652-654.

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. *Intelligence, 30*, 261-288.

Swanson, H. L. (2008). Working memory and intelligence in children: What develops? *Journal of Educational Psychology, 100*, 581-602.

Tam, H., Jarrold, C., Baddeley, A. D., & Sabatos-Devito, M. (2010). The development of memory maintenance: Six- and 8-year-olds' use of phonological rehearsal and attentional refreshment in working memory tasks. *Manuscript submitted for publication.*

Tehan, G., Hendry, L., & Kocinski, D. (2001). Word length and phonological similarity effects in simple, complex and delayed serial recall tasks: Implications for working memory. *Memory, 9*, 333-348.

Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of memory span in adults. *Memory and Cognition, 28*, 341-348.

Towse, J. N., Hitch, G. J., & Hutton, U. (2002). On the nature of the relationship between processing activity and item retention in children. *Journal of Experimental Child Psychology, 82,* 156-184.

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

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

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

Unsworth, N., Heitz, R. P., & Parks, N. A. (2008). The importance of temporal distinctiveness for forgetting over the short term. *Psychological Science, 19*, 1078-1081.

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1012-1028.

Vergauwe, E., Barrouillet, P., & Camos, V. (in press). Do mental processes share a domain-general resource? *Psychological Science*.

Wickens, D. D. (1972). Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 191-215). Washington DC: Winston.

Author Note

Correspondence concerning this article should be addressed to Chris Jarrold, Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, UK; electronic mail: C.Jarrold@bristol.ac.uk. This research was supported by a grant from the United Kingdom Economic and Social Research Council to the first and third authors (RES-062-23-0148).

Footnotes

[1] Proponents of interference-based accounts of forgetting would take issue with the implication that interference is a domain-specific process, as clearly the concept of interference is a general one. We class this as a domain-specific explanation here because the greater similarity between potentially interfering items from the same, as opposed to different, domain as the memoranda leads to domain-specific *effects*. However, clearly the mechanism by which interference effects would occur is essentially the same across all domains.

[2] Tehan et al. (2001) termed their Brown-Peterson tasks ‘delayed span’ tasks, but it is important to emphasise that these were filled with processing, unlike the delayed span tasks employed here.

[3] There are certainly other potential explanations for why spacing the storage and processing requirements of a memory task, as in complex span, would lead to superior performance relative to a condition in which these same components of the task are blocked, as in our Brown-Peterson design. These include the fact that storage items are more temporally distinct (cf. Brown, Neath, & Chater, 2007; Crowder, 1976) in the former case, and the fact that the average delay and degree of interpolated interference between presentation and recall for storage items is greater in the Brown-Peterson task. However, these accounts would appear to predict a more marked and consistent difference between these tasks than is seen here or in related work.

[4] It might be objected that interference has been conclusively shown to be the cause of forgetting from Brown-Peterson tasks, given that clear proactive interference effects are seen in classic versions of this paradigm (Keppel & Underwood, 1962; Wickens, 1972).

However, this evidence of proactive interference simply shows that individuals struggle to discriminate to-be-remembered items *across* different trials, and is not directly relevant to the issue of why the processing *within* the task causes forgetting (cf. Crowder, 1976, p.195). The fact that individuals performed less well on our versions of the Brown-Peterson task than on the comparable delayed span tasks indicates that the imposition of processing in the task does cause forgetting. It is quite possible that cross-trial proactive interference effects in this type of paradigm arise entirely because rehearsal is prevented by the processing within that trial with the result that individuals are unable to successfully maintain the memoranda in working memory (Bunting, 2006; Cowan, Johnson, & Saults, 2005; Kane & Engle, 2000).

[5] We attempted to look at this issue directly in the current work by coding and then analysing performance by each list length separately. In all three experiments the interaction between processing group, task, and list length was non-significant, even when just the complex span and Brown-Peterson tasks were compared, $F(3, 84) \leq 1.65$, $p \geq .18$. However, it should be remembered that there were only four trials at each span level, and consequently this approach may lack the power needed to detect differences between patterns of performance across list lengths.

Table 1

Descriptive statistics for Experiment 1

	Processing group			
	Verbal		Nonverbal	
	M	SD	M	SD
Baseline 1 processing RT (ms)	1608	270	1729	206
Baseline 2 processing RT (ms)	1405	217	1413	246
Baseline 1 proportional accuracy	.947	.045	.987	.021
Baseline 2 proportional accuracy	.951	.033	.971	.024
Delayed span recall score	13.03	1.79	13.05	1.62
Complex span recall score	13.60	2.08	14.28	1.35
Brown-Peterson recall score	11.28	2.67	12.53	1.90

Table 2

Descriptive statistics for Experiment 2

	Processing group			
	Verbal		Nonverbal	
	M	SD	M	SD
Baseline 1 processing RT (ms)	1248	274	1711	443
Baseline 2 processing RT (ms)	917	244	857	76
Baseline 1 proportional accuracy	.962	.028	.980	.025
Baseline 2 proportional accuracy	.982	.024	.969	.064
Delayed span recall score	13.84	1.11	13.08	1.33
Complex span recall score	12.82	1.40	12.39	2.09
Brown-Peterson recall score	10.72	2.28	11.75	1.78

Table 3

Descriptive statistics for Experiment 3

	Processing group			
	Verbal		Nonverbal	
	M	SD	M	SD
Baseline 1 processing RT (ms)	1355	307	1747	394
Baseline 2 processing RT (ms)	1088	411	912	164
Baseline 1 proportional accuracy	.942	.051	.967	.036
Baseline 2 proportional accuracy	.958	.050	.980	.021
Delayed span recall score	12.73	1.80	12.07	1.86
Complex span recall score	11.39	2.54	12.00	1.62
Brown-Peterson recall score	9.23	2.60	10.27	1.76

Table 4

Effect of processing group, and group by episode position interactions for processing times in the complex span tasks from Experiments 2 and 3

		Group		Group by position	
Response type	List length	$F(1, 56)$	p	F*	p
First	4	1.10	.299	5.24	.002
	5	0.96	.331	1.45	.218
	6	0.01	.926	0.86	.509
	7	0.33	.566	0.36	.905
Subsequent	4	0.11	737	5.22	.002
	5	0.89	.350	1.49	.205
	6	0.05	.825	2.01	.077
	7	0.24	.626	0.31	.934

* dfs = 1, 168, list length 4; 2, 224, list length 5; 5, 280 list length 6; 6, 336, list length 7.

Table 5

Analysis of processing position effects in the complex span tasks from Experiments 2 and 3 (see Figure 6).

		Linear trend		Quadratic trend	
Response type	List length	*t*(59)	*p*	*t*(59)	*p*
First	5	5.886	< .001	0.882	.382
	6	5.269	< .001	1.210	.231
	7	8.342	< .001	4.474	< .001
Subsequent	5	4.271	< .001	1.082	.284
	6	4.647	< .001	2.507	.015
	7	4.205	< .001	2.100	.040

Figure Captions

Figure 1. Schematic representation of the tasks used in each experiment. Example from list length 4 (list lengths varied from 4 to 7). S = presentation of a storage item, P = 6s of processing. (Note, in the Brown-Peterson task the processing was not divided into separate episodes, rather these are just shown to illustrate the logic of the design. Consequently, in the example shown here participants simply experienced a single block of 24 s of processing in the Brown-Peterson task).

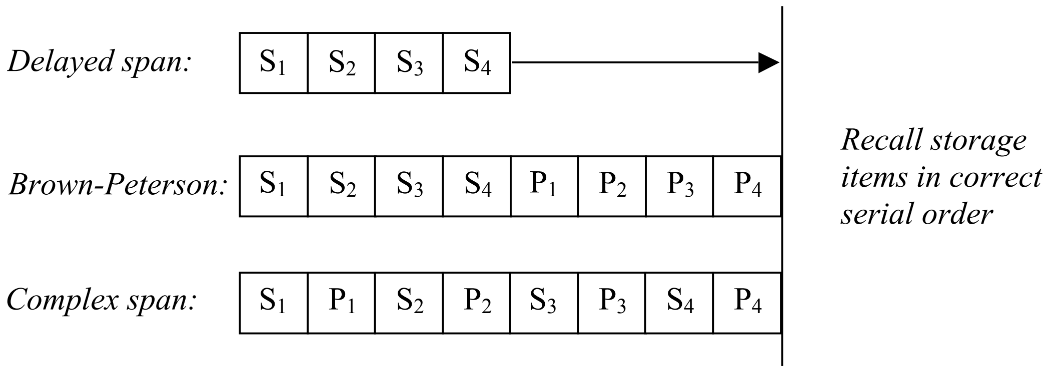
Figure 2. Effect of type of processing on complex span (CS) and Brown-Peterson (BP) performance in Experiment 1 (error bars for this and all subsequent figures are +/- 1 *SE*).

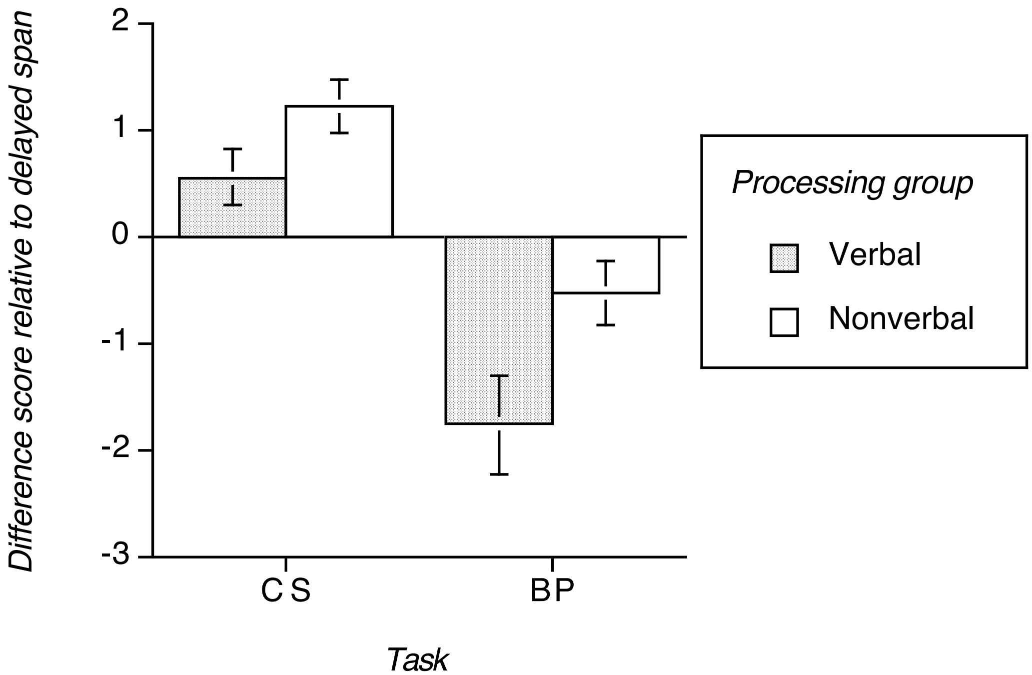
Figure 3. Effect of type of processing on complex span (CS) and Brown-Peterson (BP) performance in Experiment 2.

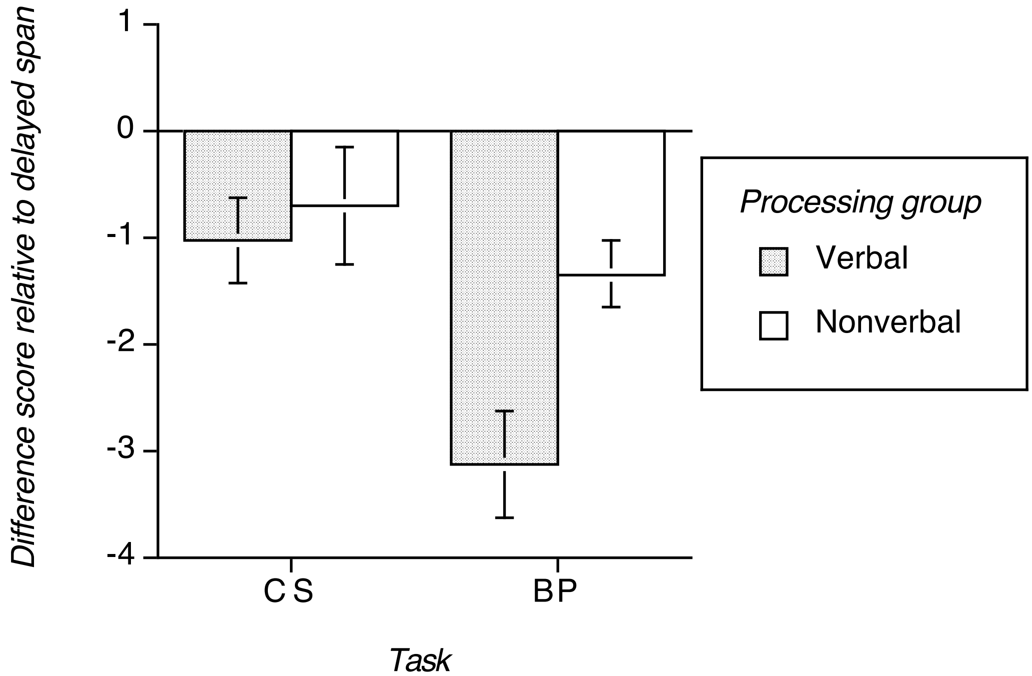
Figure 4. Effect of type of processing on complex span (CS) and Brown-Peterson (BP) performance in Experiment 3.

Figure 5. Changes in response time for either verbal (top panel) or nonverbal (bottom panel) processing across the course of the experiment among a subset of participants from Experiments 2 and 3.

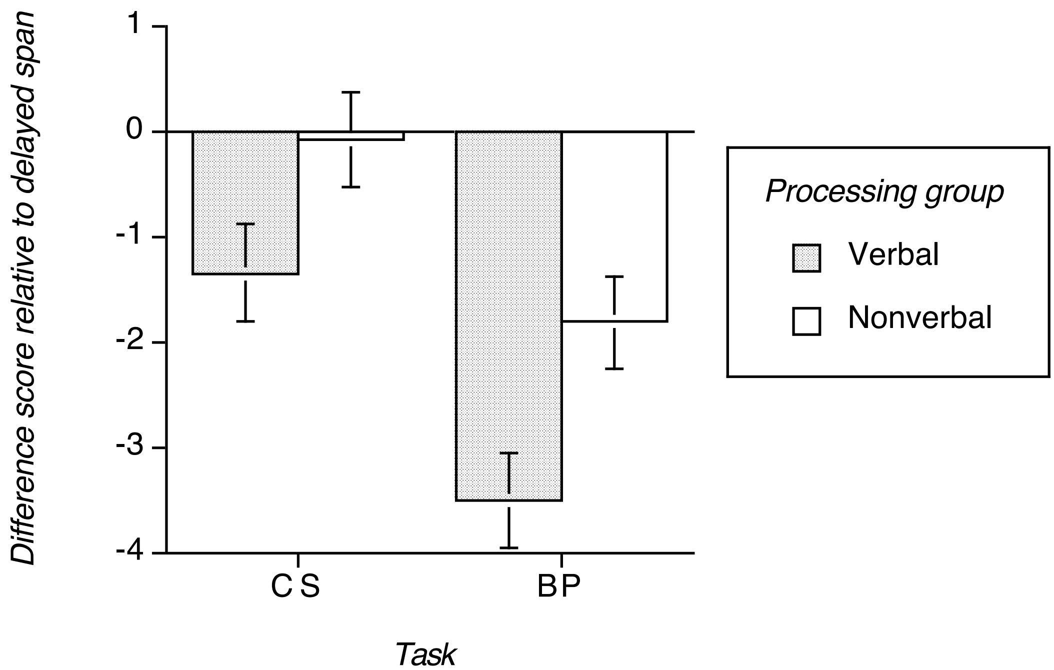
Figure 6. Changes in processing response times across the processing episodes of complex span tasks (data averaged across all participants in Experiments 2 and 3).

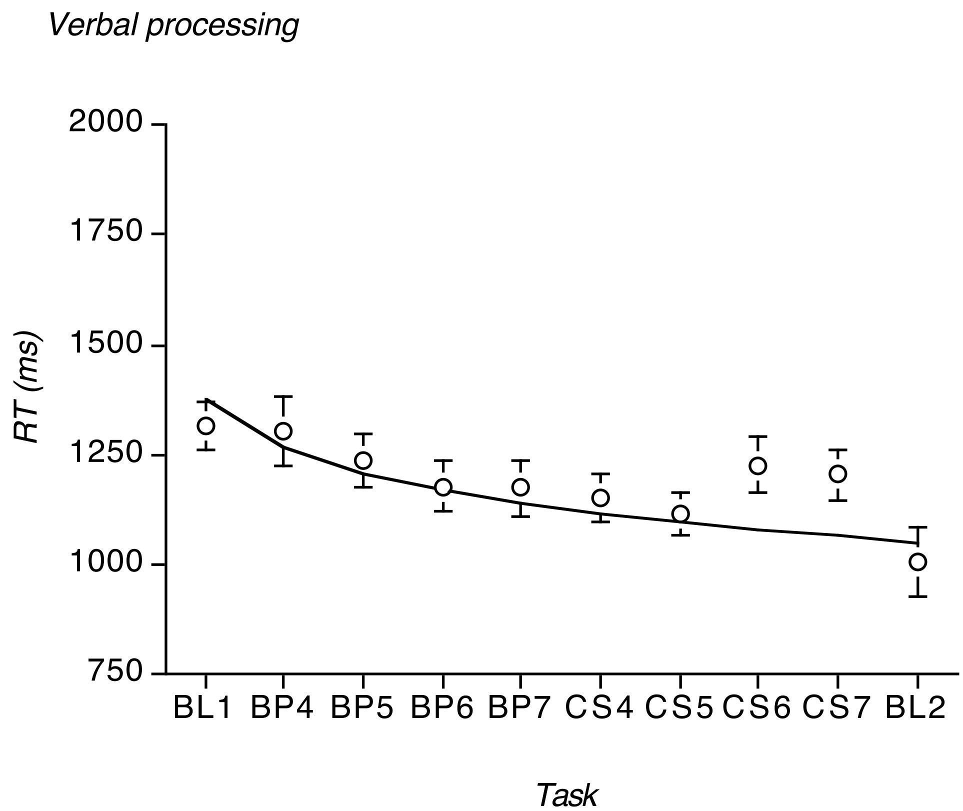

Delayed span:
S1 S2 S3 S4
Brown-Peterson:
S1 S2 S3 S4 P1 P2 P3 P4
Complex span:
S1 P1 S2 P2 S3 P3 S4 P4
Recall storage items in correct serial order

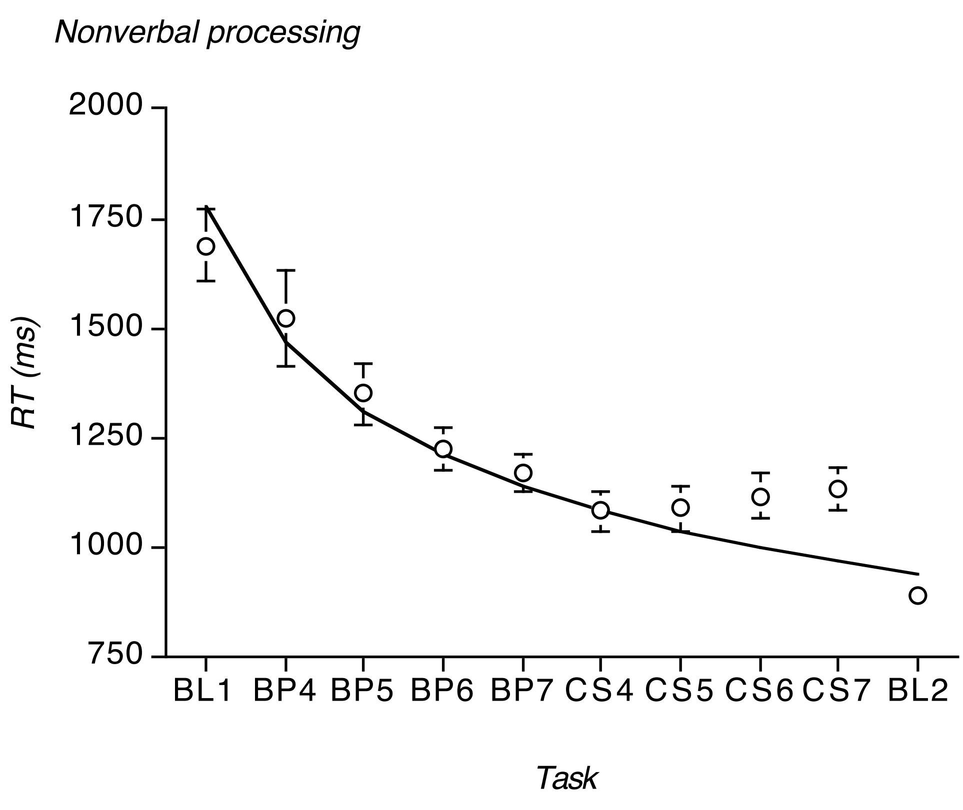

Difference score relative to delayed span
2
1
0
-1
-2
-3
C S
B P
Task
Processing group
Verbal
Nonverbal

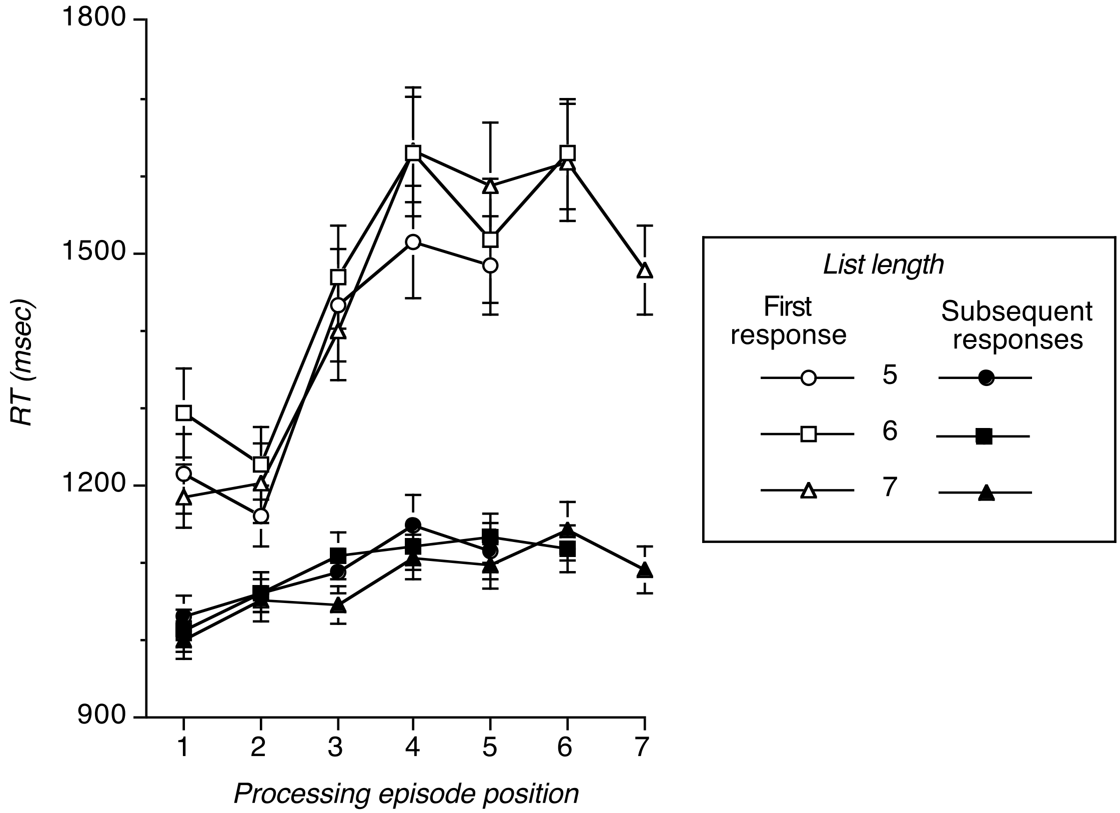


Difference score relative to delayed span
1
0
-1
-2
-3
-4
C S
B P
Task
Processing group
Verbal
Nonverbal


Difference score relative to delayed span
1
0
-1
-2
-3
-4
C S
B P
Task
Processing group
Verbal
Nonverbal


Verbal processing
2000
1750
1500
1250
1000
750
RT (ms)
BL1 BP4 BP5 BP6 BP7 CS4 CS5 CS6 CS7 BL2
Task


Nonverbal processing
2000
1750
1500
1250
1000
750
RT (ms)
BL1 BP4 BP5 BP6 BP7 CS4 CS5 CS6 CS7 BL2
Task


1800
1500
1200
900
RT (msec)
1
2
3
4
5
6
7
Processing episode position
List length
First response
Subsequent responses
5
6
7