In press, *JEP:General*

Evidence against Decay in Verbal Working Memory

Klaus Oberauer

University of Zurich

Stephan Lewandowsky

University of Western Australia

Keywords: working memory, decay, visual search, bottleneck

The research reported in this article was supported by a grant from Economic and Social Research Council UK (RES-062-23-1199) and a World University Network grant to the first author, and a Discovery Grant and Australian Professorial Fellowship from the Australian Research Council to the second author.

Correspondence should be addressed to: Klaus Oberauer, University of Zurich, Department of Psychology – Cognitive Psychology; Binzmühlestrasse 14/22, 8050 Zürich, Switzerland, Email: k.oberauer@psychologie.uzh.ch

Abstract

The article tests the assuption that forgetting in working memory for verbal materials is caused by time-based decay, using the complex-span paradigm. Participants encoded six letters for serial recall; each letter was preceded and followed by a processing period comprising four trials of difficult visual search. Processing duration, during which memory could decay, was manipulated via search set size. This manipulation increased retention interval by up to 100% without having any effect on recall accuracy. This result held with and without articulatory suppression. Two experiments using a dual-task paradigm showed that the visual-search process required central attention. Thus, even when memory maintenance by central attention and by articulatory rehearsal was prevented, a large delay had no effect on memory, contrary to the decay notion. Most previous experiments that manipulated the retention interval and the opportunity for maintenance processes in complex span have confounded these variables with time pressure during processing periods. Three further experiments identified time pressure as the variable that affected memory. The authors conclude that time-based decay does not contribute to the capacity limit of verbal working memory.

Evidence against Decay in Verbal Working Memory

People's ability to remember new information over short periods of time is severely limited. Since the early days of experimental psychology this limitation has been interpreted as reflecting the limited capacity of a short-term or working memory system (Atkinson & Shiffrin, 1968; Lewandowsky, Oberauer, & Brown, 2009a; Miller, 1956). Understanding those capacity limits is not merely of interest to theoreticians of memory: The capacity of working memory has been found to account for around 50% of the variance in general fluid intelligence (Kane, Hambrick, & Conway, 2005). Thus, the study of the capacity limits of working memory may open a window into a better understanding of a core human cognitive ability.

Most contemporary paradigms for measuring the capacity of verbal short-term or working memory test how much people can remember after a retention interval in the order of seconds (Conway, Jarrold, Kane, Miyake, & Towse, 2007; Conway et al., 2005). Because in these tasks the memoranda are presented sequentially at a pace that enables encoding of all items into working memory, any performance limitation indicative of the capacity limit of working memory must arise from rapid forgetting of part or all of the presented memoranda. Understanding what causes rapid forgetting in working memory tests would therefore illuminate what limits performance in those tasks. This would be a major step towards understanding the nature of the capacity limit of working memory.

The simplest, and for a long time very popular explanation for short-term forgetting is that memory traces decay rapidly over time (Baddeley, 1986; Brown, 1958). Recent research, however, has uncovered strong evidence against the notion of time-based decay, suggesting instead that forgetting results from some interference-based process (Berman, Jonides, & Lewis, 2009; Jonides et al., 2008; Lewandowsky, et al., 2009a; Nairne, 2002; White, 2012). Nevertheless, the issue is still hotly contended (Altmann, 2009; Barrouillet & Camos, 2009; Cowan & AuBuchon, 2008; Lewandowsky, Oberauer, & Brown, 2009b, 2009c). Here we provide further evidence against decay in verbal working memory. We limit the scope of our investigation to working memory for verbal material because there is evidence that visual representations in working memory are affected by the passage of time in a different manner than verbal representations (Ricker & Cowan, 2010). Therefore, we refrain from generalizing our conclusions to working memory for non-verbal material.

We present 10 experiments that (a) show that memory for the serial order of letters is unimpaired by the passage of time, that (b) provide independent evidence that memory restoration was prevented during the processing task, and that (c) identify time pressure as a variable that may have generated seemingly time-based forgetting in previous research.

How to Test the Decay Hypothesis

At first glance, experimentally testing the decay hypothesis seems straightforward: We need to vary the time between encoding and retrieval and test whether memory performance declines with increasing retention interval. This approach is problematic to the degree that people can engage in compensatory restoration processes such as articulatory rehearsal (Baddeley, Thomson, & Buchanan, 1975; Reitman, 1974; Schweickert & Boruff, 1986) or attentional refreshing (Barrouillet, Bernardin, & Camos, 2004; Raye, Johnson, Mitchell, Greene, & Johnson, 2007). Active restoration of compromised memory traces is commonly believed to effectively counteract decay, thus preventing its detection. An empirical test of decay must therefore prevent memory restoration during the retention interval.

There is consensus that one form of restoration, sub-vocal verbal rehearsal, can be prevented by articulatory suppression; that is, repetitive articulation of a well-known utterance such as repeating “the, the, the” or “super, super, super” (Baddeley, 1986; Page & Norris, 1998). Several studies have consistently shown that memory for verbal materials does not decline over time when the retention interval is filled with repetitive articulatory suppression (Lewandowsky, Duncan, & Brown, 2004; Lewandowsky, Geiger, Morrell, & Oberauer, 2010; Lewandowsky, Geiger, & Oberauer, 2008; Oberauer & Lewandowsky, 2008; Vallar &

Baddeley, 1982). At first glance, this provides evidence against decay because its corrosive effect failed to manifest when compensatory rehearsal was prevented. However, some theories assume that representations in working memory can be restored without covert articulation, by simply directing attention to them (Barrouillet, et al., 2004; Cowan, 2005). Therefore, a test of decay must prevent not only articulatory rehearsal but also attention-based refreshing. One goal of the present article is to present novel experimental techniques for the prevention of attentional refreshing.

Barrouillet, Camos, and their colleagues have proposed a theory that spells out the interplay of decay and attentional refreshing, the time-based resource-sharing (TBRS) theory (Barrouillet, et al., 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). The TBRS theory was developed for the complex span paradigm, in which encoding of memoranda alternates with periods of distracting activity, such as reading or mental arithmetic. From here on, we refer to those distractor tasks collectively as “processing” tasks. The core assumptions of the theory are as follows: Memory traces decay rapidly over time. Decay can be prevented by refreshing. Refreshing requires a central attentional mechanism that can carry out only one process at a time and thus acts as a bottleneck. In consequence, attention must be time-shared between refreshing and processing of distractors. By implication, the distractor task prevents refreshing for as long as it captures central attention. Because the attentional mechanism is assumed to rapidly switch between competing tasks, any brief interval of free time between two processing steps of the distractor task (e.g., between reading two words) is used for refreshing.

Thus, according to the TBRS theory, memory performance is determined by the temporal balance between the duration of attentional capture by the distractor task, during which memory traces decay, and the free time during which they can be refreshed. This balance is expressed as the *cognitive load* imposed by the distractor task. Cognitive load is defined as the proportion of time for which the distractor task captures attention. For example, if a series of arithmetic operations is required at a pace of one per second in between presentation of the memoranda, and each arithmetic operation captures attention for 0.5 s, then the cognitive load equals 0.5. If the arithmetic operations are required at a faster pace, for instance one every 0.7 s, then cognitive load increases to 0.71 (i.e., 0.5 divided by 0.7). Barrouillet, Camos, and their colleagues have repeatedly presented evidence that complex-span performance is an approximately linear declining function of cognitive load (Barrouillet, et al., 2004; Barrouillet, et al., 2007; Barrouillet, Gavens, Vergauwe, Gaillard, & Camos, 2009; Portrat, Camos, & Barrouillet, 2009).

Although initially formulated at a verbal level, the TBRS theory has recently been instantiated computationally (Oberauer & Lewandowsky, 2011). This instantiation, known as TBRS*, is the currently most precisely formulated theory about how attention-based refreshing counteracts decay, and we therefore take it as the basis for deriving predictions. In particular, we test the hypothesis that working memory is subject to time-based decay, together with the assumption that attentional refreshing, articulatory rehearsal, or both can be used to counteract that decay. Our basic approach is to independently manipulate the duration of attentional capture by a distractor task, during which decay is purportedly occurring, and the free time for refreshing, during which the corrosive effects of decay can be reversed. The TBRS theory inevitably predicts that memory should be worse when attention is engaged for a longer duration, and better when more free time is available for refreshing (see, e.g., Figure 6 in Oberauer & Lewandowsky, 2011).

Although we focus our discussion on the predictions from the TBRS theory, because it is the most precise and sophisticated formulation of a decay-restoration theory to date, we believe that our conclusions generalize to all theories assuming decay counteracted by attention-based refreshing. We return to this issue in the General Discussion.

A first methodological challenge for our approach is to find a distractor task that, while demonstrably engaging central attention, does not create representation-based interference with the memory items. If the distractor task interfered with memory representations, then extending the duration of distractor activity could increase the amount of interference and thereby lead to seemingly time-related forgetting that gives the erroneous impression of decay. For this reason, we used a visual-spatial distractor task in combination with verbal memoranda. This combination does not guarantee zero interference but is universally understood to at least minimize its contribution (Barrouillet, Portrat, Vergauwe, Diependaele, & Camos, 2011; Guerard & Tremblay, 2008; Jarrold, Tam, Baddeley, & Harvey, 2011).

A second challenge is to precisely gauge for how long a distractor task captures attention. Barrouillet and colleagues used response times for the distractor task as a proxy for the duration of attentional engagement, acknowledging that this is an imperfect estimate because not every process requires central attention. For example, it has been argued that repetition of the same distractor word (e.g., "super, super, super") does not engage the attentional bottleneck for the entire articulation duration (Barrouillet, et al., 2011). This presumed dissociation between the overt duration of an activity and its invisible underlying attentional capture presents a particularly pernicious problem because it endows the TBRS with degrees of freedom that may prevent testability: Whenever no forgetting is observed despite prolonged distraction, it could be argued that the distractor task captured attention only for a small part of its duration. Fortunately, this circularity can be broken by providing an independent empirical assay of attentional capture using a dual-task methodology.

Barrouillet et al. (2007) explicitly link the concept of central attention in their model to the notion of a central processing bottleneck as conceptualized in dual-task studies (Pashler, 1994). The notion of a central bottleneck is supported by experiments with the PRP (psychological refractory period) paradigm, which have shown that it is extremely difficult to carry out two choice tasks at the same time without mutual interference. In the PRP paradigm, people are presented two tasks, one of which ("Task 1") must be given priority. When the two choice responses are required simultaneously, or in close temporal succession, people delay the lower-priority task ("Task 2") until the bottleneck is no longer occupied by Task 1. Reviewing a large body of research on the PRP effect led Pashler (1994) to the conclusion that it is the response-selection stage in choice tasks that requires the central bottleneck. Accordingly, Barrouillet et al. (2007) showed that when a choice RT task is used as the distractor task in a complex span paradigm, memory declines as the ratio of mean RT to available time increases. In contrast, with a simple RT task, which does not require response selection, memory was unaffected by that ratio. This finding is precisely what would be predicted if the attentional bottleneck responsible for dual-task costs in the PRP paradigm is also the one that needs to be time-shared between distractor processing and refreshing in the complex span paradigm. It follows that the PRP paradigm offers a technique for determining whether, and for how long, a particular processing component in any given task requires the bottleneck. We will use this technique in the present experiments to provide independent evidence that our distractor task in the complex-span paradigm engages the attentional bottleneck (and for how long), thus enabling a powerful test of the decay assumption while preventing any speculative argument about how long the processing task diverted central attention from refreshing.

A third challenge is to devise a distractor task that enables us to manipulate the duration of processing without increasing the proportion of errors on the distractor task itself. Previous studies manipulating the time of distraction have done so by varying the difficulty of the distractor task (e.g.,Portrat, Barrouillet, & Camos, 2008). In most cases, the more difficult task variant takes longer but also engenders more errors. Errors on the distractor task, however, appear to have a detrimental effect on memory by themselves (Lewandowsky & Oberauer, 2009). Therefore, we needed to find a way to extend the duration of attentional engagement without increasing the rate of errors on the distractor task.

Overview of Experiments

We developed an experimental setting that meets the three challenges just discussed. We used a complex span task with consonants as memoranda, and a difficult visual search as distractor task in between list items. In the visual-search task, participants searched through a display of circles containing two gaps on opposite sides, except for the target, which had only one gap. Participants indicated by key press the direction of the gap (e.g., left vs. right). The duration of visual search was manipulated by the search set size (i.e., the number of circles). On some complex-span trials, all visual-search tasks used small search sets, and on other trials, all visual search tasks used large search sets. The cumulative time spent on the visual search task, and thereby the cumulative retention interval, differed substantially between trials with small and with large search sets. After each search response, a free-time period was inserted before the next stimulus. The duration of free time was varied orthogonally to the search set size, thus generating four conditions of cognitive load.

Our paradigm satisfied all three methodological challenges: First, we argue that there should be only minimal overlap between the representations needed to maintain a list of letters in order and the representations underlying visual search, because visual search involves little if any verbal representations (Zelinsky & Murphy, 2000), whereas letters are encoded into working memory primarily in a phonological format (Conrad & Hull, 1964). Therefore, on an interference account, the visual search task should not impair memory substantially. Second, we expected that increasing set size has no effect on error rate because as long as the target is not found, no response can be given, and as soon as it is found, the response is trivially easy independent of search set size. We therefore expected errors to be uncorrelated with distractor duration.

Finally, we confirmed empirically that our visual search task engaged the attentional bottleneck: Whereas there is consensus that visual search engages visual-spatial attention, it is not a priori clear whether it also engages the central attentional mechanism that acts as a bottleneck in dual-task studies, because these two forms of attention can be dissociated

(Johnston, McCann, & Remington, 1995). Our first two experiments therefore served to establish via a PRP methodology that our visual-search task actually captures central attention, and does so for longer when the search set is larger.

The remainder of the article is organized as follows: Following the demonstration that our visual search task requires central attention, we present two experiments with the complex-span paradigm, showing that the duration of attentional engagement by a visual-spatial distractor task has no impact on retention of verbal information in working memory. In one of these experiments we also prevented articulatory rehearsal though articulatory suppression. These findings provide further evidence against the presence of decay in verbal working memory, but they also raise the question why others have found that a manipulation of cognitive load with visual-spatial distractor tasks affected memory of verbal lists (Barrouillet, et al., 2007; Portrat, et al., 2008; Vergauwe, Barrouillet, & Camos, 2010). In tracking down the reasons for these discrepant outcomes, we first rule out the possibility that the discrepancy arose from the choice of distractor task by replicating our findings with the visual-spatial distractor task used by Vergauwe et al. (2010). We next note that the studies finding an effect of cognitive load with visual-spatial distractors used a method that confounded cognitive load with at least one of two other variables; the number of distractor-task errors and the degree of time pressure for carrying out the distractor task. We therefore carried out five additional experiments disentangling the effects of duration of attentional capture, distractor-task error rate, and distractor-task time pressure. To anticipate, in none of these experiments was there any effect of the duration of attentional capture on memory, contrary to the prediction based on the decay assumption. Instead, the degree of time pressure for the distractor task was identified as the variable that affected memory.

Does Visual Search Engage Central Attention?

We used the PRP paradigm to test whether, and to what extent, the search process in our visual-search task requires the central attentional bottleneck. Task 1 was an auditory-vocal choice RT task; Task 2 was the visual-search task with two levels of search set size. In addition we varied the stimulus-onset asynchrony (SOA) between the Task-1 stimulus (i.e., the tone) and the Task-2 stimulus (i.e., the search display). The logic of this design is illustrated in Figure 1, which displays two alternative theoretical scenarios. The top panel of Figure 1 shows the scheduling of processing steps under the assumption that visual search requires the central bottleneck, and therefore must wait until the central processing step of Task 1 is completed. Initial sensory processes of the search display (S) can be completed in parallel with Task 1, but the search process can only commence once Task 1 releases the central attentional mechanism; once that is the case, the search task runs uninterrupted to completion. The time for which search has to wait is independent of search set size, because set size affects only a process starting after the waiting time. Therefore, the effect of search set size should be additive with that of SOA, as indicated by the constant length of the thick unfilled arrows in the top panel.

The bottom panel of Figure 1, by contrast, shows the schedule of processing components under the assumption that visual search does not require central attention. In this scenario, visual search can proceed in parallel with central processes of Task 1. Only after visual search has finished, response selection for the search task has to wait until the bottleneck is free, because response selection is known to require central attention. Because response selection, but not search, must wait for the completion of the central processing step of Task 1, search can use that waiting time. With shorter SOA, the waiting time increases, and with it the chance that search can be completed largely or entirely within the waiting period. As a consequence, search duration has increasingly less influence on response times, and therefore, the effect of search set size becomes smaller at shorter SOAs (witness the disappearance of the thick arrow for the short SOA in the bottom panel). Therefore, this scenario predicts an underadditive interaction between SOA and search set size. At the shortest SOA (shown in the figure), the effect of set size could disappear completely.

The two scenarios illustrated in Figure 1 are extreme cases – an intermediate scenario is one where visual search requires central attention for some but not all of its duration (Oriet & Jolicoeur, 2008). The purpose our first two experiments was to gauge to what extent our visual search task required the attentional bottleneck. To that end we conducted two PRP experiments with the design of Figure 1, combining our visual search task (described below) as Task 2 with an auditory-vocal Task 1 that involved responding to one of three tones by speaking one of three arbitrarily selected words. Task 1 was designed to require difficult response selection so that it produces a long waiting time during which, if the second scenario (cf. bottom panel, Figure 1) were correct, a large part of visual search could be completed at short SOAs. Visual search tasks differ widely in their difficulty, with response-time slopes over search set size ranging from 0 to about 150 ms (Wolfe, 1998). We chose a difficult visual search task on the assumption that difficult search is more likely to engage the bottleneck than easy search. Pilot experiments revealed that our search task has a slope in the extreme upper tail of the distribution of search slopes sampled by Wolfe. Moreover, a study by He and McCarley (2010), using a visual search task very similar to ours, found that a concurrent processing task (counting backwards by threes) slows down search by increasing fixation duration on distractors. This finding increased our pre-experimental confidence that search in our task is likely to require central attention.

Experiment 1

In the first experiment we used an auditory-vocal task 1, combined with our visual search task as task 2. Search set size was manipulated with two levels; 2 versus 6 circles. SOA was varied over five levels, from 0 to 2 s.

Method

Participants. Participants were 21 students from the University of Zurich who participated in a one-hour session for partial course credit or CHF 15 (approximately US$15).

Materials. All experiments in this article were programmed in Matlab using the Psychophysics toolbox (Brainard, 1997; Pelli, 1997).

Stimuli for *Task 1* were three sine tones of 220, 440, and 880 Hz, respectively, presented for 227 ms. Participants were instructed to respond to each tone by speaking the correct associated one-syllable German word as quickly as possible. The low tone was assigned to the German word “Tisch” (table), the medium tone to the word “Post” (post), and the high tone to the word “Kamm” (comb). Response times to Task 1 were determined by automatic detection of speech onset (filtering out accidental sounds like coughing or knocking against the table).

Stimuli for *Task 2* were displays of two or six circles, scattered at random locations in an invisible 7 x 7 grid covering the whole screen. The circles’ diameter was 0.05 times the screen width. Each circle had two gaps that subtended 20 degrees of arc, one on the left and one on the right (at 9 and 3 o’clock, respectively), except for one randomly selected target circle, which had only one gap, either on the left or the right (see Figure 2 for an example display). Participants were instructed to search for the circle with only one gap, and respond as quickly as possible with the left arrow key if the gap was on the left, or with the right arrow key if the gap was on the right. RTs were measured from search-display onset to when an arrow key was pressed, whereupon the display was erased.

Design and Procedure. The design crossed two set-size conditions (2 vs. 6 circles) with 5 levels of SOA (0, 0.2, 0.5, 1, or 2 s). Within each of the 10 design cells, each of the six combinations of tones (low, medium, high) with gap orientation (left, right) was replicated 4 times, totalling 24 trials per experimental condition. The whole set of 240 trials was administered in random order.

Each trial started with a fixation cross, which was followed by one of the three tones after 0.5 s. The visual-search display followed at one of the five SOAs. Participants were instructed to respond to the tone and to the visual search task as quickly as possible, but to give the tone task priority. This priority instruction is common in PRP experiments and serves to avoid dual-task costs in Task 1. Participants first completed two blocks with 30 practice trials each, and then worked on 10 blocks of 24 test trials each.

Results

Accuracy in both tasks was high: M = .96, SD = .06 for Task 1, and M = .98, SD = .03, for Task 2. For neither task was there a significant effect of search set size or SOA on accuracy (all F's < 1.4).

Reaction times to Task 1 were analyzed for correct responses. Outliers, defined as RTs exceeding an individual's mean by three standard deviations within each design cell, were removed (1.8%). Reaction times to Task 2 were analyzed for trials on which responses to both tasks were correct; outliers (0.5%) were removed using the same criterion.

Mean reaction time to Task 1 was 825 ms (SD = 209), and it was unaffected by search set size and SOA (all F < 1), confirming that there were no dual-task costs on Task 1. Our main interest was on RTs to Task 2, which are displayed in the upper panel of Figure 3. Search set size had the expected large effect, F (1, 20) = 1092.0, p < .001, partial η^2 = .98, and there was a main effect of SOA, F (4, 80) = 74.0, p < .001, partial η^2 = .79. We used the interaction between set size and the linear contrast of SOA to evaluate the critical under-additive interaction, because this interaction best reflects the hypothesis of a monotonic decrease of the set-size effect as SOA becomes smaller. This interaction became significant, F (1, 20) = 15.4, p < .001, partial η^2 = .44. Nevertheless, there were large effects of set size at all levels of SOA, as shown in Table 1. Even at the shortest SOA, the set-size effect was reduced by only 25% of what it was at the longest SOA.

Discussion

We observed an underadditive interaction of search set size with SOA, indicative of some degree of parallelism between visual search and the central component of Task 1. At the same time, the set size effect by no means vanished at the shortest SOA. These results are in good agreement with those of Oriet and Jolicoeur (2008), who used the same experimental paradigm with a different visual-search task. One plausible interpretation of this pattern, proposed by Oriet and Jolicoeur, is that a small part of the visual search process can run in parallel with another central process, but the larger part of visual search has to wait until response selection of Task 1 has finished. Therefore, only about one quarter of the visual search duration can be completed in the waiting time imposed by the central process of Task 1. In the context of Figure 1, the results point to a situation more similar to that shown in the top panel than that in the bottom panel: About 75% of each of the shaded bars for Task 2 (including, crucially, VS) must await completion of the response-selection stage for Task 1 at short SOA's.

We need to consider one alternative interpretation: Even if visual search can occur entirely in parallel with another central process, the set-size effect reflecting search duration can be diminished at the shortest SOA only by as much as the waiting time imposed by central processes of Task 1 (see bottom panel of Figure 1). For example, suppose that the response selection stage for Task 1 (shaded portion of top-most bar in each panel) were half the duration shown in the figure, then not all of the search time for the larger set size (VS=6) could occur during the central stage of Task 1 and a small set-size effect would persist at short SOAs notwithstanding complete parallelism. Figure 4 illustrates this scenario.

This alternative scenario can be ruled out by the fact that we found a robust PRP effect for the larger set sizes in Experiment 1. The scenario in Figure 4 implies that there is no waiting time anywhere in the processing sequence of Task 2 with the large set size. Because the PRP effect (i.e., the increase of Task-2 RTs with shorter SOAs) arises from the waiting time, this scenario implies that there should be no PRP effect for Task 2 with the large set size, contrary to our finding.

Experiment 2 is a conceptual replication of Experiment 1 and in addition provides a further test of the alternative explanation sketched in Figure 4. If that scenario were correct, then a set-size manipulation with smaller set sizes should lead to a more pronounced interaction of set size with SOA, because search durations for smaller sets are less likely to exceed the waiting time imposed by the bottleneck. That is, in Figure 4, the thick arrow for the SOA of zero should disappear entirely if the VS=6 portion of the bar were shortened. In Experiment 2, we contrasted set sizes of two and three circles (corresponding to the VS=6 bar in Figure 4 to have only half its current length). If visual search could occur entirely without central attention, the set-size effect should disappear entirely at the shortest SOA in Experiment 2. In contrast, if only a small part of visual search can proceed without central attention, as suggested by Experiment 1, the set-size effect should at best be slightly diminished.

Experiment 2

Method

Experiment 2 was identical to Experiment 1 in all regards except that set-size 6 was replaced by set-size 3. Participants were 20 students from the University of Zurich.

Results

Accuracy was again high for both tasks (Task 1: M = .96, SD = .06; Task 2: M = 0.98, SD = .02). Neither set size nor SOA had an effect on accuracies of any task (all $F < 1.3$).

Reaction times were treated in the same way as in Experiment 1. The only significant effect on the RTs of Task 1 was a main effect of set size, $F(1, 19) = 6.8$, $p = .02$, partial $\eta^2 = .26$. People responded slower to Task 1 when the search set was larger (820 vs. 800 ms). There was no trace of an interaction of set size with SOA ($F < 1$).

The reaction times to Task 2 are plotted in the lower panel of Figure 3. Despite the much reduced manipulation of set size, it still had a large effect, $F(1, 19) = 353$, $p < .001$, partial $\eta^2 = .95$. The PRP effect, assessed by the linear contrast of SOA, was also significant, $F(1, 19) = 116$, $p < .001$, partial $\eta^2 = .86$. The critical interaction of set size with the linear contrast of SOA was not significant, $F(1, 19) = 2.7$, $p = .12$, partial $\eta^2 = .13$. Table 1 shows that, again, the set-size effect was substantial and significant at all levels of SOA. Relative to the longest SOA, the set size effect was reduced by at best 18% at any shorter SOA.

Discussion

The finding of a large and significant set-size effect even at the shortest SOA rules out the possibility that visual search can be carried out entirely without central attention. If visual search did not need central attention, the set-size effect of about 200 ms observed at the longest SOA should have disappeared at the shortest SOA.

The alternative scenario sketched in Figure 4 was that the waiting time enforced by Task 1 was long enough to absorb some but not all the search time at large search set sizes. This scenario could explain the results of Experiment 1 by assuming that the waiting time allows for search through about three objects in parallel with Task 1 when the SOA is zero. The remaining objects to be searched after the waiting time would then be 0 for the small set size, and 3 for the large set size. The set-size effect at the shortest SOA would then be 3-0 = 3, which is 75% of the set-size effect measured at the longest SOA (i.e., 6-2 = 4).

With these assumptions, the set-size effect should entirely disappear at the shortest SOA in Experiment 2, because search for the larger set size (of 3) should also be completely absorbed into the waiting time. This was not observed. Instead, we found a persistent set size effect at every SOA and we failed to find any interaction.

Our findings leave only one interpretation: The visual search process in our task demands central attention to a large extent. The underadditive interaction in Experiment 1, which is also visible, though not significant, in Experiment 2, shows that some part of visual search can occur during central processes of Task 1. This could be a sub-stage of search; for instance a visual analysis of the display that, once completed, enables visual attention to move to individual circles efficiently. Based on the proportional reduction of the set size effect between the longest and the shortest SOA, we estimate that the sub-stage of visual search that does not need central attention takes at most 25% of the total duration of the visual search process. We conclude that visual search in our task occupies central attention for at least 75% of its duration. This is a conservative estimate because in Experiment 2, we observed that only 18% of the search process might not require central attention (and thus 82% did). We will use this conservative estimate later to infer the attentional-capture duration in the calculation of cognitive load in the conditions of our memory experiments.

Does Time Cause Forgetting When Attention is Engaged?

Experiments 3 and 4 used the visual-search task as the distractor task in a complex span paradigm. Having established that visual search requires central attention for at least 75% of its duration, we can now use a manipulation of search set size to vary the duration of attentional capture. According to the TBRS theory, in which time-based decay can be counteracted by attention-based refreshing, increasing the duration of attentional capture, while holding the free time between visual-search trials constant, should impair memory performance. We tested this prediction in Experiment 3 as well as in Experiment 4.

It is conceivable that decay of verbal memory representations is counteracted not only by attention-demanding processes such as refreshing, but also by articulatory rehearsal. Articulatory rehearsal has been argued to require little central attention (Naveh-Benjamin & Jonides, 1984). In the theory of Cowan (Cowan, 2005), both rehearsal and refreshing play a role in maintenance of verbal memoranda, and Barrouillet and Camos have recently augmented the TBRS by adding articulatory rehearsal as a second restoration mechanism that can operate together with attention-based refreshing (Camos, Lagner, & Barrouillet, 2009). Camos and colleagues assume that the effects of rehearsal and refreshing on memory are additive, such that preventing each of them incurs measurable forgetting. On this view, time-based forgetting should be observed even when one of the two modes of rehearsal remains operative.

However, decay theorists could alternatively assume that each of the two restoration mechanisms alone is sufficient to fully counteract decay, such that preventing only one of them leaves memory performance unimpaired. As a consequence, extending the duration of visual search should not affect memory because memory could still be preserved by articulatory rehearsal. This line of argument would leave unexplained why cognitive load manipulations with non-verbal distractors have been shown to affect memory for verbal lists while articulatory rehearsal was not prevented (Barrouillet, et al., 2007; Portrat, et al., 2008; Vergauwe, et al., 2010). Perhaps because it is at odds with extant data, no current decay view instantiates this fully compensatory interaction between the two modes of rehearsal.

Nevertheless, we decided to directly test the possibility that articulatory rehearsal during visual search might completely prevent decay. Therefore, in Experiment 4 we asked participants to engage in articulatory suppression throughout the presentation phase of each complex span trial. When participants carry out a visual search task while repeating a word aloud, they cannot use central attention for refreshing, and they cannot use subvocal articulation for rehearsal. Therefore, even a decay model equipped with two fully compensatory restoration mechanisms must predict that, in Experiment 4, more forgetting occurs with longer visual-search durations.

Experiments 3 and 4

Method

The method for both experiments was identical except for the articulatory suppression requirement in Experiment 4, and the search set sizes. In Experiment 3, we used set sizes of 1 and 4, and in Experiment 4, set sizes were 2 and 6.

Participants. Twenty-six students of University of Zurich participated in Experiment 3, and 24 took part in Experiment 4. They served in a one-hour session in exchange for partial course credit or 15 CHF. One participant was excluded from analysis in Experiment 4 because that person seemed to have understood the instructions poorly, based on bad command of German, and exhibited a high error rate in the visual search task during the first 10 trials.

Materials and Procedure. The experimental task was a complex span task in which presentation of letters for later recall alternated with processing episodes in which participants completed four visual-search trials per episode. From now on we refer to these episodes of successive distractor operations between two memory items as *bursts* (see Figure 5). Memory lists consisted of 6 consonants drawn at random without replacement from all consonants except Q and Y. Materials for the visual search task were as described in the context of Experiment 1.

The design crossed two variables that together determine cognitive load as defined within the TBRS theory: The duration of attentional capture was varied by the set size of the visual-search trials, and the free time was varied through the response-stimulus interval (RSI) following each visual-search trial (200 or 800 ms). Search set size and free time remained constant within complex-span trials. There were 40 complex-span trials, 10 in each design cell, which were presented in random order. Test trials were preceded by 4 practice trials.

Each complex-span trial commenced with the presentation of a fixation cross for 2 s, followed by the first visual-search display. After participants responded, the next visual-search display followed after the RSI. After the RSI of the fourth search trial, the first letter was presented centrally on the screen. The letter disappeared after 1 s and was replaced immediately by the first visual-search display of the following distractor burst. This sequence continued until after the RSI of the fourth visual-search trial following the last letter. At that point, a question mark appeared in the center of the screen, probing for the first letter to be recalled via the keyboard (see Figure 5 for a schematic illustration). Each typed letter was presented on the screen for 200 ms and then replaced by the question mark prompting the next letter. Participants were instructed to recall the letters in order of presentation.

In Experiment 4, participants were instructed to continuously say aloud “super” during the whole presentation phase of each complex-span trial at a rate of one utterance per second. Each trial commenced with two successive displays of the word “super” centrally on the screen for 900 ms, followed by a 100 ms blank screen; the first visual-search display followed 100 ms after offset of the second “super”. Participants were instructed to read these words aloud and continue repeating them at the same pace until they saw “Stop Super” after they completed the last search trial after presentation of the last letter. The signal to stop saying “super” was replaced after 1 s by the first question mark prompting recall. Participants’ speech was recorded through a highly visible desktop microphone, and spot-checked for compliance.

Results

We first present results pertaining to response times on the visual search task, then calculate from these times the delay of recall and the cognitive load imposed by distractor processing, and finally report results for memory performance.

Visual Search. Visual search accuracy was again very high (Experiment 3: M = .99, SD = .01; Experiment 4: M = .98, SD = .02).[1] RTs from erroneous trials were excluded from analysis. For the analysis of RTs we distinguished between the first distractor burst, which precedes the memory list and therefore is unaffected by memory load (called the pre-burst from here on), and the intra-list distractor bursts following each memory item. Within the intra-list bursts, we further distinguished between the first trial in each burst, which potentially carries the task-switch cost for switching from memory encoding to visual search, and the remaining three trials within each burst (see Figure 5). Accordingly, we computed separate outlier criteria (i.e., 3 intra-individual SDs above the person’s mean) for these three sets of trials. RTs smaller than 150 ms were also regarded as outliers. Outlier RTs (0.9 and 1.3%, respectively, for the two experiments) were removed.

We analysed RTs from each experiment with two ANOVAs, one comparing pre-burst RTs (excluding the very first distractor) to no-switch intra-list RTs to assess the effect of memory load, and the other comparing no-switch and switch intra-list RTs to assess task-switch costs. Each ANOVA also included search set size and free time as variables. The statistics are summarized in Table 2. Figure 6 shows the RTs from both experiments as a function of search set size, separately for the pre-burst trials, the no-switch intra-list trials, and the switch intra-list trials.

The results can be summarized as follows. In both experiments there was a large effect of set size, which approximately doubled RTs. Visual search was somewhat slower during than before memory encoding. There were substantial switch costs, as reflected in longer RTs for the first search trial in each burst compared to the following three search trials. There were several additional significant effects that were inconsistent across the two experiments; they are of minor theoretical interest and we therefore do not consider them further.

Delay of Recall and Cognitive Load. We next computed the cumulative delays imposed by visual search with small and large set sizes by summing the mean RTs of all intra-list bursts. As shown in Table 3, increasing the search set size approximately doubled the delay filled with distracting visual search. Any theory assuming that decay is a major cause of forgetting in working memory should predict a substantial amount of forgetting to result from such a large manipulation of the retention interval.

To test the specific predictions of the TBRS theory requires computation of cognitive load; that is, the proportion of time in between presentation of memoranda during which the processing task captured central attention. Unlike most previous studies in this paradigm, our Experiments 1 and 2 provided us with an independent assay of the duration of attentional capture, thereby enabling us to compute each individual subject’s cognitive load in each design cell with unprecedented precision.

In contrast to most previous research, we were able to compute cognitive load by estimating the exact duration of attentional capture on the basis of our PRP experiments as follows. In a first step we obtained the slope of the set-size function for each individual from their average intra-list visual-search RTs. Next we determined the non-search component of the RTs in the search task (i.e., components such as response selection and execution). In

Experiment 3 we took RT at set size 1 as indicator for the non-search time (mean = 627 ms) because with a single stimulus no search for the target was required. In Experiment 4, where we did not measure times for set size 1, we computed the non-search time by subtracting the slope from the RT for set size 2, resulting in a mean non-search time of 773 ms.

We next computed the visual-search duration for each combination of set size and free time as the mean RT in that condition, minus the corresponding non-search time. Based on our PRP experiments we used the conservative estimate that visual search captures attention for about 75% of its duration; therefore we multiplied the estimated search duration by 0.75.[2]

The non-search times include response selection, which requires the bottleneck, as well as sensory and motor processes that do not. To obtain the attentional-capture duration of the non-search times, we compared the estimates to typical simple RTs, which involve no response selection. Because our estimates (627 and 773 ms) were roughly double the duration of typical simple RTs, we assumed that attention was engaged during half of the non-search time (i.e., for response selection) and thus we added half the non-search time to our estimate of attentional capture by search.

To obtain cognitive load, we then divided this total duration of attentional capture by the total processing time available per search trial in each design cell. The total time was the mean RT plus the free time and is thus unambiguously known. The procedure is illustrated in Figure 7 (top panel), and Table 4 provides a step-by-step example of the calculations with the data from Experiment 4. The cognitive load estimates thus obtained can be found in the top two rows of Table 5.

The above calculations are based on the assumption that search RTs reflect only processes serving the search task. The fact that RTs from intra-list bursts were somewhat longer than pre-burst RTs raises the possibility that search RTs reflect, in part, processes devoted to memory restoration. Therefore, we also computed a more conservative estimate of cognitive load, using the same procedure as before but estimating the duration of attentional capture during search from the pre-burst RTs only (again excluding the first trial in each pre-burst). Pre-burst RTs cannot reflect any memorial restoration processes because there is nothing to restore before the first memory item is presented. The bottom panel of Figure 7 illustrates the computation of the conservative cognitive-load estimates; the resulting estimates are presented in parentheses in Table 5.

It must be noted that the absolute values of cognitive load are of lesser interest than their variation across conditions. Table 5 confirms that our manipulations were successful, because irrespective of which estimate is used, cognitive load more than doubled between the short delay/long free time and the long delay/short free time cells. Moreover, unlike in most previous research, the values in Table 5 are derived using independently-verified estimates of attentional capture rather than imperfect proxies such as total response time.[3]

Memory Performance. Proportion of items recalled in the correct position was analyzed for each experiment with an ANOVA with search set size (2), free time (2), and serial position (6) as variables. Table 6 summarizes mean performance for the four conditions, and Figure 8 presents the serial-position curves. In Experiment 3, the main effects of both set size and free time were non-significant ($F < 1$). Their interaction just reached the conventional level of significance, $F(1, 25) = 4.2$, $p = .05$, partial $\eta^2 = .15$. Mean accuracy was .92 ($SD = .07$); the interaction reflected a slightly higher accuracy (.93) for small set sizes at long free times, and for large set sizes at short free times. The serial position curve was characterized by a strong primacy effect, together with a more confined recency effect. These trends were statistically confirmed by significant linear and quadratic contrasts, $F(1, 25) = 20.4$ (partial $\eta^2 = .45$) and 8.6 (partial $\eta^2 = .26$), respectively. None of these trends entered into a significant interaction with set size or free time.

In Experiment 4, mean accuracy was much reduced ($M = .68$, $SD = .22$), as would be expected with articulatory suppression. Yet, both the main effect of set size and the main effect of free time were non-significant ($F < 1$), as was the interaction ($F = 1.1$). There were again significant linear and quadratic trends of serial position, F (1, 22) = 55.4 (partial η^2 = .71) and 19.7 (partial η^2 = .47), respectively. The only significant interaction was between the linear contrast of serial position and free time, F (1, 22) = 5.5, p = .03, partial η^2 = .20. The primacy effect was steeper with short than with long free time.

Figure 9 shows mean proportion correct as a function of cognitive load. It is clear that memory does not decline over a substantial range of cognitive-load values, contrary to the predictions of the TBRS theory. Recall that those cognitive-load estimates were based on an independent assay of the duration of attentional capture during visual search without considering times during which the attentional bottleneck was not occupied (e.g., during half the non-search times).

Power Analysis. Could it be that we did not have sufficient statistical power to detect an effect of delay, or of cognitive load, on memory? We carried out a very conservative power analysis, based on the effect obtained by Portrat, Barrouillet, et al. (2008) with a similar design. Portrat and colleagues manipulated the duration of a distractor task, thereby increasing mean distractor RTs from 345 ms to 415 ms, an increase of just 20%; this manipulation resulted in a decrease of memory by four percentage points. If this effect reflected decay, then our manipulation—which increased delay by 80 to 100%—should generate at least twice the amount of forgetting (i.e., we conservatively assumed that quadrupling the decay time would only double forgetting). For Experiment 3, this would imply a drop in performance from 92% to 84% correct between small and large set sizes. The power to detect such an effect under the conditions of our experiment (SDs = 6.5 for small sets and 8.4 for large sets, and a correlation between these conditions of r = .78) was > .999 for a one-tailed test, computed with G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). For Experiment 4, the power to detect a commensurate drop in performance from 68% to 60% (SDs = 24 and 21, r = .93) was .995. These power calculations are based on the conservative assumption that the effect size of an 80 to 100% increase of retention interval in our experiments is only twice as large as that of a 20% increase in Portrat, Barrouillet, et al. (2008). We find no support for the proposition that our experiments lacked statistical power.

A second approach to estimating the expected effect size for our manipulations is to predict it from the difference in cognitive load, based on existing empirical slopes of memory accuracy over cognitive load. Our scoring procedure is most comparable to a study by Vergauwe, Barrouillet, and Camos (2009), who presented cognitive-load functions for an experiment in which visual-spatial materials served as memoranda and as distractors. Their data show a drop of memory performance by more than 30 percentage points as cognitive load increased from .27 to .52 (see Figure 4 in Vergauwe et al., 2009). Our manipulation of cognitive load spanned an even larger range; e.g., .21 to .54 in Experiment 3 (Table 5).[4] Power for detecting an effect of 30 percentage points under the conditions of our experiments was > .999. Even for a memory-over-load slope only one-third of that observed by Vergauwe et al. (2009), our cognitive load manipulation would still be expected to reduce memory performance by at least 10 percentage points. Power for detecting an effect of 10 percentage points was > .999 for both our experiments. We conclude that our experiments had more than sufficient power to detect an effect of decay equal in magnitude to that reported as evidence for decay in other experiments.

Computational Modelling with TBRS*. A third approach to assessing the power of our data is by comparison to simulations of TBRS* (Oberauer & Lewandowsky, 2011). We applied TBRS* to the present experiments, using the parameter values which we have shown to be best suited for TBRS* to account for a large range of data. To apply the model to the present experiment we needed estimates for the duration of attentional capture, during which memory representations decay, and of free time, during which memory is refreshed. We estimated attentional-capture duration in the same way as above when calculating the conservative estimates of cognitive load. Free time was the total available time for each trial minus the attentional-capture duration. The simulation results are shown in Table 6 alongside the experimental data.

Clearly, the model predicts a large effect of search set size on memory under the conditions of the present experiments, even larger than our conservative estimates we made above for the predicted effect size. At the same time, TBRS* predicted a fairly modest effect of free time. This is because the model predicts the beneficial effect of free time to level off (Oberauer & Lewandowsky, 2011, Figure 6). Increasing free time following each distractor beyond one second does not incur much further benefit on memory because by that time, refreshing has boosted the strength of all memory traces that are still retrievable to their maximum. We conclude that the absence of a free-time effect in our experiments is compatible with TBRS*, but the absence of an effect of search set size is not.

Discussion of Experiments 3 and 4

In two complex-span experiments, we approximately doubled the delay between encoding and recall by manipulating the duration of a distractor process that demonstrably captures central attention. In Experiment 4, we additionally prevented articulatory rehearsal. Thus, we created optimal conditions for decay to manifest itself in forgetting. Nevertheless, we observed not even a hint of an effect of the delay manipulation on memory. Complementing the absence of any effect of increased duration of attentional capture on memory, there was also a conspicuous absence of an effect of free time. Together, these findings imply that cognitive load had no effect on memory.

One potential criticism of this null finding could be that memory performance in Experiment 3 was close to ceiling, so that there was not much room for an effect to emerge. This objection can be rebutted on three grounds. First, dismissing the results of Experiment 3 on the grounds of a suspected ceiling effect would be a methodological exercise missing the theoretical point: Any viable decay theory of working memory must predict that serial recall of a six-letter list is far from ceiling after 15 to 20 s of additional delay without opportunity to refresh memory. Decay theories predicting nearly perfect performance after such delay must assume a decay rate so small that it would have only a negligible effect in standard experimental paradigms for studying working memory, thereby compromising the theory's ability to handle those other effects.

Second, the high level of performance in Experiment 3 did not prevent the effect of serial position to become clearly significant. Thus, there is evidence for forgetting of items close to the end of the list relative to items at the beginning, although our results show that this forgetting has nothing to do with the passage of time.

Third, and most important, performance was much lower and clearly far off the ceiling in Experiment 4, and yet there was no hint of an effect of temporal delay or of free time, and hence, no effect of cognitive load. Further experiments presented later in this article replicate this result several times with performance level safely below ceiling.

The substantial reduction of memory by adding articulatory suppression could be interpreted as the result of preventing articulatory rehearsal. This interpretation would imply that articulatory rehearsal contributed to memory performance in Experiment 3, and indeed this is one possibility: Even though memory representations do not decay, articulatory rehearsal might improve memory by gradually strengthening the representations of the memoranda. An alternative interpretation is that articulatory suppression adds irrelevant phonological and articulatory representations to working memory, and those representations interfere with the memoranda. The comparison of the serial position curves between Experiments 3 and 4 closely mirrors the contrast of serial-position curves without and with articulatory suppression in a previous series of experiments; those results were fit better by a model attributing the effect of articulatory suppression to interference than a model attributing it to decay and the prevention of rehearsal (Oberauer & Lewandowsky, 2008).

Why was there no cognitive load effect on memory in our experiments when strong cognitive-load effects have been reliably found in other studies? In general, the effect of cognitive load can be explained by the balance of two opposing forces acting on memory strength, a detrimental force arising from distractor processing and a beneficial force operating during free time. In decay-based theories such as the TBRS theory, decay is the detrimental force, and refreshing is the beneficial force counteracting decay. Within alternative interference-based theories such as the SOB-CS model (Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, in press), interference between memory representations and representations used for the distractor task impairs memory, and free time is used to repair the damage done by interference, for instance by removing distractor representations from memory. Lower cognitive load implies more free time for these restoration processes, whatever they are, which should translate into better memory. Most experiments showing an effect of cognitive load used distractor tasks that not only delayed retrieval but arguably also created interference with memory items. The cognitive-load effect in those experiments thus could have reflected the beneficial effect of any process that counteracts the effect of distractor interference, such as removing distractor representations, during free-time intervals.

However, free time can be beneficial only if there is a need for such measures. As long as there is no decay, and little or no interference from the distractor task, there is nothing to restore or repair. In our Experiment 3, the distractor task was not expected to interfere with memory, and we showed that there was no effect of decay either. Therefore, the distractor task did not impair memory, and consequently we found no effect of free time. In Experiment 4, which added articulatory suppression, the situation is slightly different: Whereas the visual-search task did not interfere with memory representations of letters, the repeated speaking of “super” certainly did, as was shown in other experiments (Lewandowsky, et al., 2004; Oberauer & Lewandowsky, 2008). Within the interference-based SOB-CS model, free time could be used to remove the interfering information from working memory. However, any trace of “super” removed is immediately added again because people continue saying “super” during the free time periods. Therefore, free time would not be expected to have a beneficial effect on memory in Experiment 4 either.

One question left to answer is why others have found an effect of cognitive load even when the memoranda were verbal and the distractor task was visual-spatial, such that interference between them should be minimal. For instance, Vergauwe et al. (2010) showed that memory for letters declined with increasing cognitive load when the distractor task consisted of several trials of a spatial fit judgment. We next discuss three possible explanations for this finding, and proceed to test them experimentally.

What Causes Disruption of Verbal Memory by Visual-Spatial Distractor Tasks?

We considered three hypotheses for why others have found that a visual-spatial distractor task impaired memory for verbal lists as a function of cognitive load, whereas we found no such effect. First, we considered that a difference in the distractor tasks used might be responsible. Experiment 5, reported below, ruled out this possibility by repeating the design of Experiment 3 with a distractor task that has previously been shown to generate a cognitive-load effect with verbal memoranda.

The second hypothesis is that in previous studies cognitive load was confounded with the proportion of errors on the distractor task which have been suggested to affect memory (Lewandowsky & Oberauer, 2009). We address this hypothesis through Experiments 5 to 7.

The third hypothesis arises from differences in how researchers manipulated cognitive load. Most previous studies varied cognitive load by allowing a fixed window of time during which each distractor operation could be carried out. Increasing cognitive load implied decreasing this time window, or increasing task difficulty while holding the time window constant (Barrouillet, et al., 2007; Vergauwe, et al., 2010). This procedure leads to an inevitable confound of cognitive load with the degree of time pressure for carrying out the distractor task. This confound was absent in our Experiments 3 and 4 because the distractor task was participant-paced. In Experiments 8 to 10 we de-confound time pressure and cognitive load. To foreshadow the outcome, we found no evidence that errors on the processing task affect memory. In contrast, time pressure, when manipulated independently of cognitive load, impairs memory.

Experiment 5

We repeated the design of Experiment 3 with the visual-spatial distractor task of Vergauwe et al. (2010).

Method

Participants were 21 students at University of Zurich. Two of them were excluded from data analysis because their performance on the distractor task was at chance.

The memory materials and the procedure were identical to Experiment 3, but we exchanged the visual search distractor task for the spatial-fit judgment task of Vergauwe et al. (2010). Participants had to judge whether a bar fitted into a gap between two small rectangles arranged horizontally (see Figure 10); the bar fitted in a random half of trials. We manipulated the difficulty of the judgment by varying the degree to which the horizontal extension of the bar differed from the size of the gap (2 vs. 20 pixels difference). On half of the complex-span trials, participants received only hard fit-judgment trials, and on the other half, they received only easy fit-judgment trials. This manipulation served to manipulate the attentional-capture duration in a manner analogous to the set size manipulation in Experiments 3 and 4.

Results and Discussion

Distractor Task. Easy distractor trials were responded to with 98% accuracy (SD = 1.8), and hard trials with 69% (SD = 14). Response times were treated and analysed as in Experiments 3 and 4; the results of the two ANOVAs are reported in Table 2. As expected, RTs were substantially larger for difficult than for easy fit judgments: mean RTs (excluding pre-bursts) were .72 s (SD = .13) for easy, and 1.12 s (SD = .38) for hard trials. The difficulty manipulation increased RTs by 50% (see Table 3)—not as much as the manipulation of search set size in the previous two experiments, but more than the difficulty manipulation of

Portrat, Barrouillet, et al. (2008). Increasing difficulty added a substantial cumulative delay of 9.5 s to the retention interval of the first memory item. The only other significant effect was the switch cost: RTs were slower for the first trial after encoding a letter than for subsequent trials. Notably, intra-list RTs were not slower than pre-burst RTs, rendering it unlikely that people strategically withheld their responses during intra-list distractor trials.

To conclude, we successfully varied the time required for the spatial fit judgment in the distractor task used by Vergauwe et al. (2010). Unlike the manipulation of set size in visual search, the present manipulation targeted the difficulty of response selection, which is a processing component known to require central attention. Therefore, in agreement with Vergauwe, et al. (2010), we assumed that the more difficult judgment condition involved a longer duration of attentional capture. To compute cognitive load, we estimated the duration of attentional capture as the mean distractor RT in each condition, minus 0.3 s for non-central processes (this estimate is based on typical simple RTs; e.g., Barrouillet, et al., 2007). The summed attentional-capture durations in a processing burst were then divided by the total time of that burst, which is the sum of the RTs and the free-time intervals. The resulting cognitive-load estimates are presented in the bottom row of Table 5.

It is possible that the switch cost in fit-judgment RTs reflect strategic postponement of responding to the first fit judgment to gain time for refreshing. We therefore again calculated a conservative estimate of cognitive load based on the assumption that the switch cost reflected additional free time for refreshing (values in parentheses in Table 5). It is clear that both estimates of cognitive load vary substantially over the four experimental conditions.

Memory performance. Proportion correct in the four conditions is presented in Table 6. There was no main effect of distractor-task difficulty, no main effect of free time, and no interaction (all $F < 1$). The only significant effect was that of serial position, $F(5, 90) = 9.5$, $p < .001$, partial $\eta^2 = .34$. The serial-position curves in this experiment and all those that follow had the same shape as those in the preceding two experiments, and therefore we do not present them again. The serial-position effect shows that there were statistically detectable effects on memory. However, as in the preceding two experiments, memory did not depend on our manipulations of temporal variables. Our manipulation of cognitive load was as strong as in the preceding two experiments, so based on published effects of cognitive load we should again expect a drop in memory accuracy of 20 to 30 percentage points. Power to detect a difference of just 10 percentage points was > .999 (one-tailed test, SD = 7.2 and 5.7 for the easy and hard condition, respectively, and $r = .38$).

Table 6 also includes the results of a simulation of Experiment 5 with TBRS*, which generated a predicted a drop in memory due to delay of 17 percentage points. Clearly, this prediction is at odds with the data. For Experiment 5, TBRS* also predicted a sizeable effect of the free-time on memory (i.e., a drop of 10 percentage points), contrary to the data.

We conclude that we created a substantial manipulation of cognitive load with the same distractor task as Vergauwe et al. (2010), and yet this manipulation had no effect on memory, contrary to the predictions of the TBRS theory, and of decay theories in general. The discrepancy in outcome between Vergauwe et al. and our results therefore cannot be due to differences in the distractor tasks.

Experiments 6 and 7

The results of Experiment 5 already cast doubt on the assumption that the proportion of errors in the distractor task affects memory performance, because people made many more errors in the difficult than the easy distractor condition but memory was unaffected by distractor difficulty. This finding runs counter to our earlier work, which revealed distractor-task errors as a predictor of memory performance in a regression analysis of the data of Portrat et al. (Lewandowsky & Oberauer, 2009). To obtain more robust data on this issue, we carried out two more experiments to investigate the role of errors on the distractor task more thoroughly. Both experiments again used visual search as the distractor task, but instead of manipulating search set size we varied the difficulty of the decision about the location of the gap in the target by using an oblique decision boundary (oriented along the NW – SE axis). In the difficult condition, the gap was very close to the boundary, whereas in the easy condition it was about 90 degrees away from the boundary. Experiment 7 differed from Experiment 6 only in that we added error feedback to each distractor trial to ensure that people noticed their errors, thus maximizing the impact of potential post-error processing.

Method

Participants. Twenty-two students from University of Zurich participated in Experiment 6, and 21 participated in Experiment 7. Two participants in Experiment 6, and one in Experiment 7, were excluded from analysis because their memory performance was perfect; this leaves N=20 for each experiment.

Materials and Procedure. Materials and procedure were the same as for Experiment 3, with the following exceptions. The manipulation of search set size was replaced by a manipulation of decision difficulty in the search task and the search set always comprised two objects. Thus, free time and decision difficulty varied between complex-span trials. People were instructed to detect the target circle with a single gap and decide on which side of a boundary the gap was. The boundary was an imaginary line running through the center of the circle at an angle of 45 degrees. The boundary was illustrated by a figure in the written instructions. The upper-left side of the boundary was assigned to the left key and the lower-right side was assigned to the right key. The location of each gap was drawn from a uniform distribution of locations within a pre-determined range of angular distances from the boundary. For easy trials, the distances from the boundary ranged from 80 to 90 degrees, and for difficult trials, the distances ranged from 0 to 45 degrees.

In Experiment 7, each response was followed by visual feedback: The target circle turned green for a correct response, or red for an error. The display was cleared 0.2 s later. A RSI of 0.2 or 0.8 s followed each response (Experiment 6) or each feedback display (Experiment 7) to manipulate free time.

Results and Discussion.

Accuracy and RTs for the distractor task as well as memory performance are summarized in Table 7. The manipulation of distractor difficulty had the desired effect. In Experiment 6, distractor accuracy was 98.7% (SD = 1.5) for easy, and 78.8% (SD = 12.5) for hard trials. In Experiment 7, the corresponding values were 98.0% (SD = 1.4) and 83.3% (7.5). In Experiment 6, RTs were 1.03 s (SD = 0.20) for easy and 1.31 s (SD = 0.30) for hard trials; in Experiment 7, RTs were 0.98 s (SD = 0.14) and 1.29 s (SD = 0.20), respectively. Increasing the decision difficulty increased the retention interval by about one third (see Table 3).

Memory was unaffected by decision difficulty ($F < 2.1$ in both experiments) and by free time ($F < 1$). The interaction of these two variables reached significance in Experiment 7, $F(1, 19) = 4.9$, $p = .04$, partial $\eta^2 = .21$, but not in Experiment 6 ($F < 1$). The only strong effect in both experiments was that of serial position, $F(5, 95) = 10.8$, $p < .001$, partial $\eta^2 = .36$ in Experiment 6, and $F(5, 95) = 13.7$, $p < .001$, partial $\eta^2 = .42$ in Experiment 7. As shown above, the manipulation of distractor decision difficulty in Experiments 6 and 7 affected RTs, most likely through an effect on response selection. Therefore, the more difficult distractors engaged central attention for a longer time. Yet again, a substantial increase of the duration of attentional capture had no effect on memory.

Moreover, in three experiments (Experiments 5 – 7) we consistently failed to find an effect of the proportion of distractor-task errors on memory. This result goes against the outcome of our re-analysis (Lewandowsky & Oberauer, 2009) of the data from Portrat, Barrouillet, et al. (2008). In the Portrat, Barrouillet, et al. experiment, attentional-capture duration was confounded with proportion of distractor-task errors, and in our multi-level regression analysis only the latter predicted memory. Experiments 5 to 7, by contrast, show that distractor-task errors per se do not disrupt memory.

In light of these results it remains unclear what caused the decline in memory by 4 percentage points in the experiment of Portrat, Barrouillet, et al. (2008). We observed no hint of an effect of attentional-capture duration on memory across five experiments, each with a stronger manipulation of this duration than the mere 20% time increment of Portrat and colleagues. We obtained no evidence for an effect of distractor-task errors across three experiments, even though our distractor error rates varied more than those of Portrat and colleagues. We can only speculate that an as yet unidentified difference between the distractor task of Portrat and colleagues (i.e., judging whether a stimulus is above or below the screen midline) and the two distractor tasks used here might be responsible for the sole aberrant finding of Portrat, Barrouillet, et al. (2008). Clearly, however, given the outcome of our 5 experiments, the effect observed by Portrat et al. was not due to the passage of time per se.

Having ruled out an unknown idiosyncrasy of our distractors (by Experiment 5) and having ruled out post-error processing (by Experiment 6 and 7) as the explanation for the cognitive-load effect in extant precedents, we turn to an examination of our final hypothesis: Does the cognitive load effect with visual distractors and verbal stimuli arise from differences in time pressure?

Time Pressure and Cognitive Load

As noted earlier, most published experiments manipulating cognitive load allowed a fixed time window for each operation of the distractor task, and higher load was created by either decreasing the available time while holding processing time constant, or conversely, by increasing processing duration while holding the total available time constant (Barrouillet, et al., 2004; Barrouillet, et al., 2007). This procedure inevitably confounds cognitive load with the degree of time pressure because, whichever way cognitive load is manipulated, it reduces the available time relative to the time needed. Time pressure has been shown to affect cognitive processes in non-trivial ways in decision making and reasoning (Evans, Handley, &

Bacon, 2009; Svenson & Edland, 1987), rendering an effect of time pressure on working memory plausible.

The following three experiments served to de-confound time pressure and cognitive load within our visual-search distractors from Experiments 1-4. The key to doing this is to manipulate independently the time available for carrying out each visual search trial (which determines the degree of time pressure) and the free time added after processing of the search task has terminated (see Figure 11). In the three experiments reported below, we imposed a deadline on each visual-search trial, thereby controlling the available time for search. The search display was erased when the deadline was reached, or when the person responded, whichever came first. Display offset was followed by an experimenter-controlled free-time period. Time pressure was manipulated through the placement of the deadline relative to the time required for the visual-search task (estimated as described in the context of each experiment). Free time was manipulated independently.

With this method we can vary time pressure while holding cognitive load largely constant. Cognitive load is defined as the ratio of attentional-capture duration to total time. When a deadline is imposed to generate time pressure, attentional capture by the search task is shortened by the deadline, because visual search cannot continue after the display is erased, curtailing processing on a subset of trials. Note, however, that if people found the target before reaching the deadline, they can still complete response selection and execution after the deadline. Thus, attention-demanding processing components will on occasion reach briefly into the free-time period after the deadline. As time-outs will occur more often with high time pressure, post-deadline completion of processing will, on average, result in a modest shortening of free time with high compared to low time pressure. Because greater time pressure shortens attentional-capture time as well as free time, cognitive load can be expected to remain roughly constant as time pressure is increased. Effectively, time pressure squeezes all time components, thereby leaving their ratio largely unchanged. The relationship between the different time parameters under varying levels of time pressure is illustrated schematically in Figure 11.

We note that there is no way to hold cognitive load perfectly constant when time pressure is varied, because at least one of the determinants of cognitive load, attentional-capture duration, cannot be experimentally controlled but only measured. This is true even when every experimenter-controlled feature of the task is held constant, because attentional-capture duration depends on what people *do*. Cognitive load is thus never an experimenter-controlled construct: In all relevant research to date, cognitive load has been *measured* rather than being an experimenter-determined level of an independent variable. Ultimately, therefore, the success of our effort to vary time pressure and cognitive load independently can only be evaluated by measuring cognitive load. In Experiment 10 we will present a detailed measurement of cognitive load under two levels of time pressure and confirm that our method enabled independent variation of both variables.

Building on the method illustrated in Figure 11, we carried out three experiments varying cognitive load independently of time pressure. In Experiments 8 and 9 we manipulated cognitive load by varying the experimenter-controlled free time; in Experiment 10 we varied cognitive load through a variation of search set size.

Experiment 8

In Experiment 8 we used the method illustrated in Figure 11 to independently vary time pressure and free time. Our main aim was to investigate the effect of time pressure on memory independent of cognitive load.

Method

Participants. Nineteen students from the University of Zurich participated in the experiment in exchange for 15 CHF or course credit.

Materials and Procedure. Materials and procedure were largely the same as for Experiment 3: Participants memorized lists of six consonants, and each consonant was followed by four trials of the visual-search task, with a constant search set size of four objects. The major difference from the preceding experiments was that we imposed a deadline for presentation of the search display. The search display disappeared whenever a response was entered or the deadline was exceeded. Because visual search cannot continue in the absence of the search display, the deadline places a time limit on search (though not necessarily on response execution). We selected a deadline of 1.2 s for high time pressure, and 1.9 s for low time pressure. These times were taken from the distribution of RTs for set size four in Experiment 3. In that experiment, 51% of trials were completed after 1.2 s, and 90% of trials were completed after 1.9 s. Free time was manipulated orthogonally to time pressure: Erasure of the search display was followed by either 0.2 s or 0.8 s of RSI with a blank screen.

Our manipulation of time pressure differs in an important way from the variation of time pressure implied by the common manipulation of cognitive load. In conventional cognitive-load manipulations people are given a constant time window to complete a distractor operation. The deadline on each operation coincides with the onset of the next distractor, such that time pressure is inextricably linked to the total time between two distractors. Hence any increase in time pressure inevitably implies an increase of cognitive load. Here we decoupled the deadline from the onset time of the next distractor, and thus we can vary time pressure and total time independently. Specifically, we added a pre-determined free-time interval after people’s responses, or after the deadline, whichever came first. In this way we decoupled the manipulation of time pressure from the variation of cognitive load.

Results and Discussion

Time pressure had the expected effect on visual-search performance. At low pressure, participants committed on average 5.5% time-out errors (SD = 5.0), whereas at high pressure, they committed 47.0% time-out errors (SD = 9.7). Those proportions mesh well with the data of Experiment 3. Of the trials responded to in time (timely RT’s from here on), 95.3% were correct with low time pressure (SD = 2.6), compared to 87.2% (SD = 6.5) with high pressure.

Timely RTs were analysed as before by two ANOVAs, summarized in Table 8. The first ANOVA showed that pre-burst RTs were faster than intra-list RTs, and time pressure resulted in faster RTs. Thus, under time pressure people increased speed at the expense of accuracy. The effect of time pressure was more pronounced for intra-list RT's, and it was more pronounced when free time was short. These effects are plotted in the top panel of Figure 12. In addition, in replication of all preceding experiments, the second ANOVA revealed a significant task-switch cost.

In sum, the analyses of RTs under time pressure show the same trends as in the preceding experiments with self-paced distractor tasks, although attenuated by the time limit (as reflected in the smaller scale of Figure 12 compared to the preceding RT plots). The switch cost was also observed in the proportion of time-out errors (analyses not shown). This finding shows that the slowing of RTs on switch trials does not just reflect strategic postponement of responses to allow for additional refreshing, because in the present experiment, postponing the response does not buy any additional time. It is more plausible that the effects of switching and of list position arise from variables outside the person's intentional control, such as the time for reconfiguring the task set.

As shown in Table 9, memory accuracy was worse under high than under low time pressure, $F(1, 18) = 10.4$, $p = .005$, partial $\eta^2 = .37$, and worse with short than with long free time, $F(1, 18) = 15.0$, $p = .001$, partial $\eta^2 = .46$. In addition, there was the usual main effect of serial position, $F(5, 90) = 26.0$, $p < .001$, partial $\eta^2 = .59$. All interactions were non-significant ($F < 1.5$).

To conclude, time pressure, manipulated independently of duration of attentional capture and free time, had an adverse effect on memory. The present experiment thus reveals one source of the cognitive-load effect obtained in so many previous experiments. Most of those experiments confounded cognitive load with time pressure, and we have shown that time pressure alone causes forgetting. In addition, in the presence of time pressure, free time has a beneficial effect, which was not observed in the experiments above using a self-paced distractor task. We carried out Experiment 9 to replicate the effect of time pressure and to test a hypothesis about the origin of the free-time effect.

Experiment 9

One potential cause of the beneficial effect of free time under conditions of time pressure is the following: When encoding of each letter is preceded and followed by distractor activity under time pressure, there might not be sufficient time to fully encode and consolidate each letter into WM. Because it takes time to switch from visual search to memory encoding, and from memory encoding back to visual search, the actual time available for encoding might be less than the nominal 1 s presentation time of letters. Longer free-time intervals preceding each letter might have ameliorated this problem in Experiment 8 because the free-time period could already be used to switch to the encoding task, so that more of the presentation duration of the letter can be used for actual encoding. If this is the case, extra time immediately after each letter should have the same effect.

To test this incomplete-encoding hypothesis, in Experiment 9 on half the trials we gave people extra time for encoding each letter. We again limited the time available for visual search. Instead of manipulating free time after each distractor, on half the complex-span trials, we inserted a 1 s time gap between each letter and the first visual-search trial (with the remaining half of trials containing no gap). If people had difficulty fully encoding or consolidating the memory items in Experiment 8, then their memory should benefit from this additional encoding time.

Method

Participants. Participants were 106 students from the University of Western Australia who took part for course credit. Four of them returned incomplete data. Of the remaining 102 participants, 27 were excluded because their accuracy of visual search fell below 70% in the low time-pressure condition.

Materials and Procedure. Materials and procedure were as in Experiment 8, with the following differences: Instead of varying the free time after each search trial, we manipulated the presence or absence of a temporal gap of 1 s between the presentation of a letter and the onset of the first visual search display. Unlike all experimental manipulations in the preceding experiments, we varied the gap between blocks of 12 trials. This was done to make the gap even more salient for participants, thereby encouraging its use for further consolidation. Order of blocks was counterbalanced across participants. Time pressure was varied independently of gap in the same way as in Experiment 8 (i.e., available time for visual search was 1.2 or 1.9 s). The search display disappeared whenever people responded or the deadline was reached; this was followed by a blank-screen RSI of 0.2 s. Because of time constraints, participants completed only 6 complex-span trials per time-pressure condition in each block. Both blocks were preceded by two additional practice trials for a total of 24 trials.

Results and Discussion

The time-pressure manipulation was again effective. Participants committed 34.2% time-out errors in high-pressure trials (SD = 14.0), compared to 5.9% in low-pressure trials (SD = 15.7). Of the trials completed in time, 78% (SD = 10.0) were correct in the high-pressure condition, and 89.2% (SD = 7.0) in the low-pressure condition.

The two ANOVAs carried out on timely distractor RTs are summarized in Table 8, and mean RTs are plotted in the lower panel of Figure 12. People responded faster when time pressure was high, and this effect was larger for intra-list RTs. Thus, the manipulation of time-pressure again shifted people's speed-accuracy trade-off criterion in favour of speed. There were again switch costs, reflected in slower RTs for the first search trial in each processing burst compared to the following three search trials. The switch cost was much reduced in the condition with a temporal gap. Apparently people used the gap for carrying out whatever process was responsible for the switch cost in the condition without a gap.

Memory performance is summarized in Table 9. Time pressure had a detrimental effect on memory, $F(1, 74) = 27.6$, $p < .001$, partial $\eta^2 = .27$. The presence of a temporal gap had no effect, and the two variables did not interact (both $F < 1$). There was again a main effect of serial position, $F(5, 370) = 90.7$, $< .001$, partial $\eta^2 = .55$, which did not interact with any other variable (all $F < 1$).

The absence of any effect on memory of an added 1 s gap between presentation of each memory item and the onset of the first visual-search display disconfirms one possible explanation for the free-time effect in Experiment 8: Free time is not used to complete or consolidate encoding into WM. This result also has an important implication for the possibility that the switch cost (observed in the no-gap condition and in all previous experiments) reflects postponement of the first distractor operation for squeezing in additional refreshing immediately after encoding. Switch costs in the previous experiments were always less than 1 s: Given that a 1-second gap had no beneficial effect in the present study, any time gained by postponement in the previous studies equally would not have resulted in any measurable benefit for memory.

Experiments 8 and 9 showed that increasing time pressure creates a detrimental effect on memory that previous experiments might have misidentified as an effect of cognitive load because those studies confounded cognitive load with time pressure. When those two variables are deconfounded, it is time pressure not cognitive load that adversely affects memory (Experiment 8).

However, so far we only deconfounded one part of the cognitive-load equation, namely free time. To complete our reanalysis of the cognitive load phenomenon, we must also deconfound the second component, namely the duration of attentional capture. This was the purpose of our final experiment.

Experiment 10

The final experiment in this series independently manipulated time pressure and duration of attentional capture. We again used visual search as a distractor task; duration of attentional capture was manipulated through search set size as in Experiments 3 and 4, and time pressure was manipulated via the duration of the search display. Figure 13 illustrates the effect of these manipulations. Unlike in Experiments 8 and 9, time pressure was calibrated for each participant individually to ensure that subjective pressure was equal across participants.

Method

Participants. Two groups of students from the University of Zurich were enrolled in this experiment (N = 24 and N = 21, respectively). The only difference between the two groups was that for the second group we measured RTs also for responses given after the deadline (i.e., during the RSI). One participant from the first group was excluded because of perfect memory performance; one participant from the second group was excluded because of excessively long visual-search durations in the calibration phase (see below).

Materials and Procedure. The complex-span trials of the present experiment followed the general procedure of the preceding experiments. Participants recalled lists of six consonants, presentation of each of which was preceded and followed by four visual-search trials. We independently varied the search set size (2 vs. 6 objects) and the degree of time pressure. Free time following each visual-search trial was held constant at RSI = 0.5 s. Time pressure was individually calibrated for each participant as described below. Individual calibration was necessary in this experiment to ensure that the time-pressure conditions were comparable for large and small search set sizes which differed considerably in mean duration.

The experiment commenced with a calibration phase, consisting of four blocks of visual-search trials, half at each set size. Half the participants received set sizes in order 2, 6, 6, 2, and the other half received the order 6, 2, 2, 6. Search displays were presented for a limited time window, which was initially set to 1.5 s and 2.2 s for the small and the large set size, respectively. In each block, these time windows were adapted by the weighted up-down algorithm (Kaernbach, 1991), an efficient psychophysical algorithm for calibrating testing conditions to a desired proportion of successful responses. We initially calibrated the presentation time for visual search at both set sizes to 20% time-out errors while holding accuracy > 90%. This was accomplished by decreasing presentation time of the search display by 50 ms every time a correct response was made before the deadline, and increasing presentation time by 200 ms every time people missed the deadline (i.e., a time-out error). Erroneous responses before the deadline elicited error feedback (i.e., display of the German word for "error" for 300 ms), but did not lead to adaptation of presentation time. Each block continued until 20 successive trials were completed with an average accuracy of 90% or higher (not counting time-out errors).

After the fourth block, the adapted presentation times were averaged across the two blocks for each set size; these adapted times reflected the individual's time demand for a given set size. For each set size, an individual's presentation time for high time pressure was set to 0.6 of the mean adapted time, and the presentation time for low time pressure was set to 1.2 times the mean adapted time. By this procedure, the degree of time pressure was expressed as a constant proportion of an individual's time demand for achieving the same level of performance at a given set size. Thus, we manipulated time pressure in essentially the same way as in the preceding experiments, by setting presentation times to two different proportions of people's estimated time demand for a given set size. In that way we also equated time pressure across set sizes because for each level of time pressure, the same proportion of the separately-calibrated times was used for both set sizes.

In the main phase of the experiment, participants completed 48 complex-span trials (i.e., 12 in each design cell) in random order. Except for the temporal parameters, trials were identical to those in Experiment 8.

Results

Effect of time pressure on visual search. Calibrated presentation times for small search sets were 0.63 s (SD across subjects = 0.23) for high time pressure and 1.26 s (SD = 0.47) for low time pressure. Presentation times for large sets were 1.29 s (SD = 0.25) for high pressure, and 2.59 s (SD = 0.50) for low pressure. Recording of responses after the deadline enabled a more detailed analysis of the effect of time pressure, which served as a manipulation check and to gain insight into how time pressure affects processing in the context of a working-memory task. This analysis is detailed in the Appendix.

To summarize, our analysis revealed that the offset of the search display at the deadline did not immediately stop processing; rather, processing continued for some time into the following free-time period. Moreover, higher time pressure did not increase processing efficiency, but rather led people to sacrifice accuracy for speed. The incurred loss in accuracy was larger for larger set sizes. Therefore, if anything, our manipulation of time pressure affected search through large sets more than search through smaller sets. This rules out the possibility that the effect of time pressure on memory might have been weaker with larger set sizes, thereby obscuring a hypothetical effect of decay.

Delay of Recall and Cognitive Load. The analysis in the Appendix has shown that distractor processing does not immediately stop when the search display is erased. Therefore, we used the RTs of all responses, including those in the RSI following the deadline, as the basis for calculating cognitive load. These times were only available for the second group of participants; therefore our estimates are based on that sub-sample. For the small set of distractor trials on which participants did not respond even until after the free time, we set RT to the time elapsed during the search trial (i.e., presentation time plus free time), assuming that participants continued processing until the onset of the next search display.

With low time pressure, the cumulative delay imposed by the search task (i.e., the sum of search RTs within each complex-span trial) was 18.5 s (SD = 2.0) with small set sizes, and

34.4 s (SD = 5.2) with large set sizes, such that the set-size manipulation added 15.9 s of delay (85.9%). Under high time pressure, cumulative delays were 16.8 s (SD = 1.9) for small set sizes, and 28.3 s (SD = 5.0) for large set sizes, implying that 11.5 s of delay was added by the set-size manipulation (68.4%).

The duration of attentional capture was computed in the same way as for Experiments 3 and 4: We first determined each person's search slope and non-search time for each condition of time pressure. Search time was then calculated by subtracting non-search times from RTs. Note that those estimates were not tied to the physical deadline because people could respond after the deadline (i.e., after the stimulus disappeared). Attentional capture duration was again taken to be 0.75 times the search time, plus 0.5 times the non-search time.

Total time for each distractor trial was obtained by adding the free time (0.5 s) to the time of offset of the display—which was the lesser of the RT or the deadline. The resulting estimates of cognitive load can be found in Table 10. It is clear that search set size had a substantial effect on cognitive load, whereas time pressure had none, exactly as intended.

Memory performance. Time pressure again had a detrimental effect on memory, F (1, 42) = 16.9, p < .001, partial η^2 = .29. Search set size had no measurable effect despite increasing retention interval by more than 10 seconds for the first item, F (1, 42) = 1.5, p = .24, partial η^2 = .03. The familiar effect of serial position was again significant, F (5, 210) = 47.7, p < .001, partial η^2 = .53. None of the interactions reached significance (all F < 1.5). The fact that additional processing time did not impair memory when time pressure was held constant provides strong evidence against the notion of decay because in this paradigm, withholding of responses would not have gained additional time for memory restoration.

Recall accuracy is presented as a function of cognitive load in Figure 14, showing that time pressure impairs memory whereas cognitive load has no statistically detectable effect. It is important to remember that time pressure was calibrated separately for each level of cognitive load (i.e., set size), thereby ensuring that high and low time pressures, respectively, were operationalized identically between levels of cognitive load.

Based on the accuracy-over-load slopes of Vergauwe et al. (2009) we can expect that an increase of cognitive load by .13, as induced by our manipulation of attentional-capture duration (see Table 10), should lead to a drop of memory performance by about 10 percentage points. Power to detect such an effect under the circumstances of our experiment (i.e., SDs of 11 and 12 for small and large set sizes, respectively, and r = .83) was virtually 1.0. Power to detect an effect of just half that size was still .998.

Testing an interference account of the time-pressure effect. Why should distractor-task time-pressure affect memory? From the perspective of interference models such as SOB, the detrimental effect of distractor processing on memory arises because representations involved in distractor processing interfere with representations of the memoranda (Oberauer & Lewandowsky, 2008). Carrying out a task under high time pressure arguably involves more cognitive control than under low time pressure. Among other things, the cognitive system must estimate how much time is left until the deadline, and if the deadline has passed, it must try to stop a belated response to avoid responding at a time when the next trial’s display is already on. Stopping a response in the midst of its preparation has been studied extensively with the stop-signal task (Logan & Cowan, 1984). Cognitive control relies on representations (e.g., representations of the duration of the available time window, or of the goal to stop a response to avoid spill-over into the next trial) which potentially interfere with representations of memoranda. For instance, Verbruggen and Logan (2008) have shown that the goal to stop a response is automatically associated to the stimulus. This raises the possibility that stopping goals, or other representations involved in cognitive control, are encoded into working memory and thereby interfere with list items.

We explored this possibility, drawing on a prediction from one interference-based model, SOB (Farrell & Lewandowsky, 2002; Oberauer & Lewandowsky, 2008). In SOB, encoding depends on the novelty of events relative to what is already encoded in memory in the same context (e.g., in the same serial position). Applied to the complex-span paradigm, this means that consecutive distractors following a memory item add to interference to the degree that they differ from each other. In contrast, when the same distractor is processed several times in a row within the same processing burst (i.e., following the same item), interference does not increase with the number of repetitions. We confirmed this prediction in several earlier experiments using word reading as distractors: Each memory item was followed by reading of a single word, three different words, or four repetitions of the same word. Interference was the same for a single word and four repetitions of the same word but was considerably greater when three different words were read (Lewandowsky, et al., 2010). Here we apply this successful prediction of SOB to interference from representations involved in cognitive control. We assume that every search trial in which people miss the deadline (i.e., time-out errors) triggers control processes, such as an attempt to stop the response before the onset of the next search display. We further assume that when several time-out errors occur within the same processing burst (i.e., following the same letter), the representations involved in control are highly similar and therefore do not add much interference over and above a single time-out error in this processing burst.

To test this assumption, we regressed recall accuracy on each trial on four predictors in a multi-level model (Pinheiro & Bates, 2000): (1) search set size, (2) time pressure, (3) number of processing bursts with at least one time-out error, and (4) number of additional time-out errors over and above the first time-out error in each burst. The first two predictors dummy-coded the experimental conditions; based on the ANOVAs above we expect that time pressure but not set size affects recall. If the first time-out error in each processing burst creates interference but further time-outs in the same burst do not add further interference—as expected by SOB—the third predictor (number of processing bursts with at least one time out) but not the fourth (number of additional timeouts) should contribute to predicting recall performance.

We estimated linear mixed-effects (LME) models using the function *lmer* in the lme4 package (Bates, Maechler, & Boker, 2011) for R (R-Development-Core-Team, 2011). The model sought to predict the number of items recalled in the correct position in a trial, using a binomial link function. In addition to the fixed effects of the four predictors we included random effects (i.e., individual differences between subjects) on the intercept and the effects of set size and time pressure. Starting from the full model we successively removed predictors to test whether a simpler model achieved a better fit, as evaluated by the Bayesian Information Criterion (BIC). As predicted, the best-fitting model retained only time pressure and number of processing events with at least one time-out as predictors. Crucially, the number of additional timeouts failed to predict memory performance. Table 11 summarizes parameter estimates from the full model and the final model.

Discussion

The results of Experiment 10 provide another dissociation of cognitive load and memory performance. When cognitive load was manipulated through a variation of attentional-capture duration, and time pressure was varied independently, only time pressure, but not attentional-capture duration, had an effect on memory. When considered together with Experiment 8, which provided a similar dissociation between time pressure and free time, the overall pattern of results suggest the conclusion that the apparent cognitive-load effect observed in many previous experiments did not arise from “cognitive load”—as defined by the ratio of attentional capture to total processing time—but rather from time pressure, which to date has always been inextricably linked with cognitive load.

Our regression analysis further illuminates the effect of time pressure: Holding experimental conditions constant, recall was worse the more items were followed by at least one time-out error in the following processing burst. The total number of time-out errors within each processing burst, in contrast, was unrelated to memory. This result is as predicted from the SOB model, together with the assumption that every time-out error creates a representation that enters working memory and interferes with the memoranda.

General Discussion

The results of our experiments provide a serious challenge for decay theories. We introduced a large manipulation of retention duration, increasing delay of recall by 25 to 100% across experiments, while demonstrably blocking attentional refreshing and (in one experiment) also articulatory rehearsal. Yet, we observed no hint of more forgetting with longer retention intervals. The present findings add to the growing evidence against decay in verbal WM (Lewandowsky, et al., 2009a). Previous work blocking articulatory rehearsal (by repeating the same word aloud) and in some instances also central attention (by choice RT tasks) has also shown little or no effect on memory of a substantial increase in retention interval (Lewandowsky, et al., 2004; Lewandowsky, et al., 2010; Lewandowsky, et al., 2008; Oberauer & Lewandowsky, 2008).

One potential loophole for decay theories that had been left open by previous work was that the duration of attentional capture by the distractor tasks was not independently established. That is, although participants were continually processing distractors in all earlier studies (Lewandowsky, et al., 2004; Lewandowsky, et al., 2010; Lewandowsky, et al., 2008; Oberauer & Lewandowsky, 2008) it was conceivable that part of the measured processing time did not occupy the attentional bottleneck, thus allowing for some attentional refreshing. Therefore, decay proponents could use assumptions about attentional capture as an unconstrained free parameter in theorizing and argue that any additional decay during prolonged delays could be fully compensated by additional time for refreshing. The present work closes this loophole.

We used the PRP paradigm to ascertain that the processing stage whose duration we manipulated requires central attention. This enabled us to calculate the duration of attentional capture, and the cognitive load imposed by it, to make unambiguous predictions about the effect that our manipulations should have on memory. According to any decay model, increasing visual search duration under the conditions of our experiments should lead to substantial additional forgetting. According to any theory assuming that attentional refreshing serves memory restoration, shortening the free time between search trials should also lead to further forgetting. These two assumptions form the core of decay-based theories assigning a role to attention in memory maintenance, most prominently represented by the TBRS theory. We additionally used a computational implementation of that theory, TBRS*, to formally derive the above predictions through simulations. The results of six experiments (i.e., Experiments 3, 4, 5, 6, 7, and 10) resoundingly falsified these predictions.

The importance of obtaining independent measurements of whether, and how long, a distractor process captures central attention is further illustrated by a recent set of experiments by Barrouillet, De Paepe, and Langerock (2012). They used verification of arithmetic equations as distractor task in a complex-span procedure and manipulated the duration of each verification judgment though the equations' format: People responded faster to equations composed of digits (e.g., "2 x 3 = 6") than to equations with number words (e.g., "two x three = six"), and recall was worse with word equations than with digit equations. The free time after each response was held constant. Barrouillet et al. (2012) assumed that word equations captured central attention for longer than digit equations, so that word equations imposed a higher cognitive load. This assumption can be tested with the PRP procedure in an analogous manner to our PRP experiments, and such a study has been done: Sigman and Dehaene (2005) found an underadditive interaction between format and SOA, such that the RT difference between word and digit format completely disappeared at the shortest SOA. This implies that whatever process is responsible for the slower responses to word equations as opposed to digit equations can run entirely in parallel with other processes. Applied to the study of Barrouillet et al. (2012), the results of Sigman and Dehaene (2005) imply that the word equations increased the total time for distractor processing (i.e., RT + free time) without increasing the duration of attentional capture. As a consequence, cognitive load was actually smaller for word equations than for digit equations. Therefore, the observed decline in memory with word equations arguably constitutes a dissociation between memory and cognitive load, similar to what we observed here.

Potential Objections

Could the decay notion be salvaged in light of our findings? One objection that might be raised by proponents of decay is that we have not prevented articulatory rehearsal in all our experiments, and therefore, participants could have fully compensated the effect of decay by articulatory rehearsal. The results of Experiment 4, in which we added articulatory suppression to the attention-demanding visual search task, already question this possibility. Although articulatory suppression was very effective in reducing memory performance relative to Experiment 3, it did not lead to any further forgetting as a function of increased processing duration, which speaks against the possibility that articulatory rehearsal could have masked the operation of decay. Here we add a further principled argument against this potential defence of decay theory.

Decay theorists assuming more than one kind of restoration process, such as articulatory rehearsal and attention-based refreshing (e.g, Camos, et al., 2009), must make assumptions about how these restoration processes interact. There are two logical possibilities, a compensatory and a non-compensatory model. The compensatory model assumes that each restoration process on its own is sufficient to fully counteract the effect of decay. Hence, when one of the two processes is prevented, the other can fully compensate for it. As a consequence, forgetting over time should be observed only if both restoration processes are prevented. A theory making this assumption could be insulated from much of the empirical evidence against decay – though not from the result of our Experiment 4 – but at the cost of losing most of its explanatory power. This is because experimental conditions that prevent both articulatory rehearsal and refreshing are rare. Most instances of forgetting from short-term or working memory that have so far been attributed to decay are routinely observed under conditions that may prevent at most one but not both restoration processes. For instance, the word-length effect is obtained under conditions that do not prevent attention-based refreshing (Lewandowsky & Oberauer, 2008); articulatory suppression impairs memory without preventing attention-based refreshing; conversely, the cognitive-load effect has been observed with visual-spatial distractor tasks that do not prevent articulatory rehearsal (Vergauwe, et al., 2010). The compensatory model cannot explain these phenomena as manifestations of decay because whichever restoration process in those experiments was blocked, the other one should have fully compensated and performance should have been unimpaired. It is not surprising, therefore, that as far as we know, no current theory subscribes to this compensatory view.

The non-compensatory model, by contrast, has been explicitly endorsed by Camos et al. (2009) as an extension to the TBRS. The additive model states that a single restoration process is insufficient to fully counteract decay. As a consequence, preventing one of the two restoration processes must be sufficient to render memory representations vulnerable to decay. Thus, blocking of one restoration process must cause some time-based forgetting albeit less than if both processes were blocked. Proponents of this model can account for the results just mentioned (e.g., the effects of articulatory suppression), but they cannot explain the absence of time-based forgetting when refreshing is prevented for a prolonged period of time, as observed consistently across six experiments in the present article. For the same reason, the non-compensatory model cannot explain the absence of time-based forgetting when articulatory rehearsal is prevented for a substantial duration, as observed in the present Experiment 4, and in about a dozen other experiments (Cowan et al., 2006; Lewandowsky, et al., 2004; Lewandowsky, et al., 2010; Oberauer & Lewandowsky, 2008; Phaf & Wolters, 1993; Vallar & Baddeley, 1982). In summary, neither version of the dual-refreshing approach can be reconciled with the existing data.

Another objection that has been raised against previous experiments is that the distractor task was self-paced, giving participants the opportunity to postpone responding to surreptitiously squeeze in some refreshing, thereby fully compensating any effect of decay during the manipulated retention interval (Barrouillet, et al., 2011). The present series of experiments addresses this objection in four ways. First, we obtained an independent measure of the duration of attentional capture of our visual-search distractor task, which enabled us to estimate the duration of attentional capture more precisely than any previous study. Second, we computed a conservative estimate of cognitive load for our self-paced experiments that assumes that *all* slow-down of RTs during the memory task—relative to a pre-list burst—reflects strategic postponement used for refreshing. This, in effect, instantiated a most extreme version of the postponement hypothesis, and even when thus correcting for any potential postponement, we found no effect of cognitive load on memory performance. Third, we tested the effectiveness of postponing responses by introducing a free-time gap of a full second after every memory item, and found no beneficial effect (Experiment 9). Fourth, in Experiment 10 we manipulated cognitive load with a deadline method, as demanded by Barrouillet et al. (2011), and still found no effect of temporal delay on memory.

In this context, we discovered that the method recommended by Barrouillet and colleagues – controlling time by providing a fixed window during which people can respond – implies a confound between cognitive load and time pressure. We showed in three experiments that time pressure, when deconfounded from free time or attentional-capture duration, has a detrimental effect on memory. Conversely, neither the duration of attentional capture (Experiment 10) nor free time (Experiment 8) affects memory when time pressure is controlled, and interference between memoranda and distractors is minimized.

We conclude that the conventional method of manipulating cognitive load is problematic because it is inevitably confounded with a variable that we have isolated as being causally related to memory performance—namely, time pressure during the distractor task. Given that virtually all examinations of cognitive load to date have confounded temporal parameters with time pressure, our results have considerable implications.

Implications for Theories of Working Memory

No role for decay. The primary implication of our work is that it renders it difficult if not impossible to salvage the notion of decay as a non-negligible cause of forgetting in verbal WM. At this juncture we cannot rule out the possibility that non-verbal information in WM, by contrast, may be subject to decay. Nevertheless, our results challenge numerous theories that appeal to temporal decay as a general mechanism by which forgetting occurs in short-term and working memory (Anderson, Reder, & Lebiere, 1996; Baddeley, 1986; Barrouillet, et al., 2004; Daily, Lovett, & Reder, 2001; Hitch, Towse, & Hutton, 2001; Kieras, Meyer, Mueller, & Seymour, 1999). Importantly, some of those theories have been instantiated computationally, which circumvents the notorious tendency of verbal models to escape testability via semantic reinterpretation (Farrell & Lewandowsky, 2010; Lewandowsky, 1993; Lewandowsky & Farrell, 2011). For instance, we showed that our own computational instantiation of the TBRS theory, TBRS* (Oberauer & Lewandowsky, 2011) cannot handle the present results. It unambiguously predicts a substantial amount of forgetting due to decay when the duration of the distractor operations is increased as much as in our experiments. For the same reason, models based on the ACT-R architecture (Anderson, et al., 1996; Daily, et al., 2001) must predict more forgetting with longer delays filled by our visual-search task, because memory restoration in ACT-R depends on firing of productions, for which the ACT-R architecture assumes a strictly serial bottleneck.

In the theory of Cowan (1995, 2005), attention serves as a store with a limited capacity of about four chunks that protects memory representations from decay. Maintenance of a list exceeding the capacity of the focus can be accomplished by articulatory rehearsal and by cycling subsets of items through the focus, thereby refreshing them. Thus, when articulatory rehearsal is prevented and the focus of attention is engaged by a distracting task, at least some items should be prone to forgetting through decay. In Cowan's theory, choice response tasks are assumed to demand attention, and our PRP experiments showed that our visual search task compete with response selection. Therefore, the visual-search task must have consumed at least part of the attentional capacity in Cowan's theory, leaving less capacity for refreshing memoranda. Thus, in the context of Cowan's theory, decay should lead to more forgetting with longer distractor-filled retention intervals, contrary to what we observed.

Finally, theories of decay in which articulatory rehearsal is the sole restoration process, most prominently represented by the phonological-loop model (Baddeley, 1986) and its computational implementations (Burgess & Hitch, 1999; Kieras, et al., 1999; Page & Norris, 1998) would also encounter difficulties accounting for our findings. These models are uniformly challenged by the absence of forgetting as a function of time filled with articulatory suppression (i.e., the present Experiment 4, as well as numerous previous experiments: Lewandowsky, et al., 2004; Lewandowsky, et al., 2010; Oberauer & Lewandowsky, 2008; Vallar & Baddeley, 1982).

Forgetting by interference. In contrast, an alternative class of theories, which explains forgetting in WM through interference between representations (Oberauer & Kliegl, 2006; Oberauer & Lewandowsky, 2008; Oberauer, et al., in press; Saito & Miyake, 2004), has no difficulty explaining our findings. The visual search task arguably has little overlap in content or process with the memorization of letters in serial order. Unless memory for the letter list involves spatial features, or the visual search task recruits verbal representations, there is no reason to expect much interference between these tasks. Carrying out visual search in between letter encoding should therefore have little, if any, effect on memory. Consistent with this prediction is the observation that recall was very good in most of our experiments, even somewhat better than what was found in other experiments testing serial recall of six consonants without any distraction (Farrell & Lewandowsky, 2003; Henson, Norris, Page, & Baddeley, 1996; Oberauer, 2003). Only when articulatory suppression was added, did performance drop substantially. This drop can readily be explained by the fact that articulatory suppression generates interfering articulatory and phonological representations.

Deconstructing “cognitive load”. The notion of decay, combined with attention-based refreshing, has received much support from experiments showing a strong decline of memory with increasing cognitive load. The present experiments rule out decay as a major cause of forgetting in verbal WM, and therefore imply that an alternative explanation needs to be found for the cognitive-load effect. We have provided relevant initial evidence.

Specifically, our Experiments 8-10 identified time pressure during distractor processing as a crucial causal variable that is routinely confounded with cognitive load: Whenever time pressure is inadvertently manipulated together with cognitive load, the “cognitive-load” effect arises, although as we have shown here, it does not reflect the balance between time for decay and time for restoration as purported by the TBRS theory. When time pressure is either absent (Experiments 3-7) or deconfounded from other temporal parameters (Experiments 8 – 10), then cognitive load has no effect on performance. We next flesh out a more detailed explanatory account of “cognitive load” within an interference approach.

From the perspective of an interference-based theory of WM, we need to distinguish two cases, within-domain and cross-domain manipulations of cognitive load. Within-domain manipulations use distractor tasks that engage representations in the same content domain as the memoranda (e.g., memory for letters combined with word reading as distractor task). In this situation, processing of the distractor material involves generating representations that interfere with representations of the memoranda. Interference is less severe when cognitive load is low because the additional free time in between individual distractor operations can be used to partially repair the damage done by preceding distractor operations. This restoration process may involve rehearsal or some other mechanism, such as removal of unwanted information from working memory. We favour the latter view in our theorizing (Oberauer, et al., in press) but those implementation details are unimportant for present purposes. One implication of this explanation, which sets it apart from the decay-based explanation proposed by Barrouillet and colleagues, is that the cognitive-load effect is primarily an effect of the amount of free time, whereas the duration of attentional capture should have only a small effect on memory at best.[5] We have confirmed this prediction in three experiments combining verbal memoranda with verbal distractors (Oberauer, et al., in press).

Cross-domain manipulations of cognitive load, which were the focus of the present experiments, combine memoranda and distractor tasks from different content domains, such as memory for verbal items combined with a visual-spatial distractor task. A cross-domain cognitive-load effect can be explained in two ways without appealing to decay. First, many memory tasks and many distractor tasks, although purely verbal or purely non-verbal by appearance alone, actually engage a mixture of both verbal and non-verbal (e.g., visual-spatial) representations. For instance, reading words involves eye movements, which are known to disrupt spatial WM contents (Lawrence, Myerson, & Abrams, 2004). Conversely, processing of spatial distractors could involve verbal self-instructions that would interfere with verbal memoranda. Thus, cross-domain interference can arise for the same reason it arises within domains; namely, by the encoding of distracting information, albeit in an attenuated manner. However, this mechanism did not seem to operate in the present memory experiments. If representations of the visual-spatial distractor material had interfered with the memoranda, free time should have enabled participants to reduce the effect of interference through restoration. Contrary to that expectation, free time had no effect in our experiments unless they included a manipulation of time pressure.

A second explanation emerged from the present Experiments 8 to 10, which showed that even when distractor representations do not interfere with the memoranda, time pressure in the distractor task is by itself sufficient to impair concurrent memory. Because the cross-domain manipulation of cognitive load in most published reports to date has been confounded with time pressure (Barrouillet, et al., 2007; Vergauwe, et al., 2010), we believe that time pressure has likely contributed to those effects. Differences in time pressure could also have contributed to the effect of arithmetic-problem format (digits vs. words) on memory reported by Barrouillet et al. (2012). In that study participants were shown each problem for a limited time, thereby creating time pressure. It is not clear whether and how the authors calibrated the presentation duration to the time demand for each problem format, so it cannot be ruled out that time pressure was more severe for the word problems than for the digit problems, and this could contribute to explaining why memory was worse with word problems.

Whereas we have established the role of time pressure in cross-domain interference, its involvement in within-domain cognitive load effects is less certain. We already noted that the balance between representational interference and restoration is sufficient to explain the within-domain cognitive-load effect. It is however possible that time pressure contributed further to the effect in those cases where cognitive load was confounded with time pressure, as in most research to date (e.g., Barrouillet, et al., 2004; Vergauwe, et al., 2009). Whether or not this is the case needs to be determined by further research that de-confounds the two variables in the way we did in Experiments 8-10.

Finally, why should distractor time pressure affect concurrent retention in WM? At present we cannot offer a definitive answer. One explanation that might come to mind is that under time pressure people shift cognitive resources away from memory maintenance and towards distractor processing in an effort to speed up processing (Morey, Cowan, Morey, & Rouder, 2011). Our data do not support this explanation: The analysis of cumulative accuracy over time in Experiment 10 (reported in the Appendix) shows that people do not carry out visual search more efficiently under high time pressure; rather, they respond to time pressure by sacrificing accuracy for speed.

From the perspective of interference theories of WM, an explanation would start from two plausible assumptions: More severe time pressure recruits increased cognitive control, and cognitive control uses representations that potentially enter WM and thereby interfere with representations of the memoranda. We can think of several representations enrolled in control processes for coping with time pressure, among them estimates of how much time is left for a response, the goal to work quickly, and when the deadline has been missed, a goal to stop responding lest the response spill over into the next trial. Our regression analysis of memory performance in Experiment 10 lends initial support for an interference-based account by demonstrating that multiple time-out errors are related to memory in the same way as multiple interfering distractors: Speaking a distracting word once after each memory item severely impairs recall, but repeating the same word several times adds no further interference over speaking it once. In the same way, committing a single time-out error after each memory item predicted a decline in recall, but additional time-out errors did not add to the predicted loss of memory. This pattern is precisely as predicted by one interference-based model, SOB (Lewandowsky, et al., 2010; Oberauer & Lewandowsky, 2008).

It remains a fruitful avenue for future research to flesh out a control theory that can help identify the processes by which time pressure causes interference in working memory. One route through which cognitive control creates interference with verbal memoranda could be the enrolment of verbal representations for control. Overt verbalization of task-relevant information has been shown to boost cognitive control in the task-switching paradigm, whereas articulatory suppression impairs control (Kirkham, Breeze, & Marí-Beffa, 2012; Kray, Eber, & Karbach, 2008). The complex-span paradigm requires task switching. It is therefore not implausible that people under time pressure recruit verbal representations to boost control processes that enable a rapid switch from memory encoding to visual search, and to ensure that a response is given before time expires. Along similar lines, research by Tubau, Hommel, and López-Moliner (2007) suggests a role for phonological codes in controlling the timing of learned action sequences. Under time pressure, accurate timing is important for anticipating the deadline to make a timely response without being unnecessarily early. Thus, there are hints in the literature for a role of verbal representations in cognitive control, which could provide a starting point for investigating how control results in interference with verbal working-memory contents.

Conclusions

We presented the first study to provide independent evidence for the involvement of an attentional bottleneck in a distractor task used in a complex-span paradigm. We presented five complex-span studies using this distractor task which showed no evidence of temporal decay despite the retention interval being increased by 25 to 100% by extending the duration of attentional capture due to the distractor activity. We de-confounded several variables that were confounded in previous experiments manipulating cognitive load. We demonstrated that time pressure during distractor processing, but not the duration of distractor processing, affects memory, casting doubt on the interpretation of published effects of cognitive load in terms of decay and refreshing.

The consistent failure to observe time-based forgetting when the attentional bottleneck was demonstrably occupied and when articulatory rehearsal was prevented presents a very strong challenge to decay models. The data instead support interference-based approaches, such as those recently instantiated in the SOB model (Oberauer & Lewandowsky, 2008; Oberauer, et al., in press).

Footnotes

(1) One participant consistently reversed the response keys but had inconspicuous response times; the responses of that person were reverse-coded, resulting in 98% correct.

(2) The 75% estimate comes with a margin of error. Fortunately, the precise proportion of visual search time that captures the bottleneck affects our manipulation of cognitive load only quantitatively. To illustrate, we calculated cognitive load for Experiment 4 on the assumption that search captures the bottleneck for only 50% of the time and on the assumption that it captures the bottleneck 100% of the time. The 50% estimates were .27, .41, .34, and .45 for the four conditions in the order of Table 5; the 100% estimates were .33, .50, .53, .70. Both estimates of cognitive load varied substantially across conditions.

(3) In contrast to our cognitive-load estimates, which might underestimate cognitive load—especially in the conservative variant—published estimates based on total response time (e.g., Vergauwe, et al., 2010) certainly overestimated cognitive load, and they do so to an unknown extent, because the actual duration of attentional capture was not measured.

(4) Although Vergauwe et al. (2009) used visual-spatial materials, Vergauwe et al. (2010) have shown that the span-over-load slopes are of comparable size for verbal memoranda and distractors as they are for visuo-spatial materials.

(5) Duration of attentional capture can still be expected to have some effect because the extent to which information is attended to is a powerful determinant of how strongly it is encoded into memory (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Phaf & Wolters, 1993). Therefore, distractors attended to for a longer time would be encoded more strongly, thus creating more interference. However, because encoding into working memory is fast, we assume that the effect of attentional-capture duration asymptotes quickly (Oberauer, et al., in press).

Appendix: Analysis of Response Time Distributions in Experiment 10

Our analysis is based on the sub-sample for which RTs were measured during the RSI. Figure A1 shows cumulative distributions of correct responses as a function of time since display onset. Each data point reflects the proportion of all trials that was completed with a correct response up to that time. These curves provide a measure of performance that integrates speed and accuracy. It is clear from this figure that in conditions with high time pressure the curves continue to rise for several 100 ms after the deadline, indicating that processing was not immediately disrupted by display offset. This is also reflected by the observation that responses given during the RSI were correct on 84% and 82.5% of trials for small and large set sizes, respectively.

We fit a descriptive model to these data: The first part of the cumulative distribution, until the deadline for display offset, was described by a negatively accelerating exponential function with three parameters: time of offset from zero (a), rate (b), and asymptote (c). This exponential growth curve continued for the second part, during the 0.5 s following the deadline, with rate br, where r is a free parameter expressing the proportional reduction in rate after display offset. At the end of the RSI, when no responses to the previous trial were recorded any more, the function turned into a flat line.

The model was fit within a mixed-effects framework (Pinheiro & Bates, 2000), which expresses parameter group means as fixed effects and individual differences in parameter values as random effects, using the *nlme* package (Pinheiro, Bates, DebRoy, & Sarkar, 2005) in R (R-Development-Core-Team, 2011). The model allowed for main effects of set size and of time pressure, as well as their interaction, on parameters a, b, and r; effects of these variables on asymptote c were not expected because with infinite time our search task is trivially easy (as reflected in the high accuracy in the self-paced experiments, and the fact that c was estimated to .999). We tested these effects by model comparison, fixing each effect to zero if that did not lead to a loss of fit, as assessed by the Bayesian Information Criterion

(BIC). Only the interaction of set size and time pressure on *b* could be fixed to zero by this criterion.

Table A1 shows a summary of the fixed-effect parameter estimates. Set size had a modest effect on the delay until the curves offset from zero (*a*) and a large effect on their growth rate. This pattern reflects the fact that the fastest correct responses were given at about the same time after display onset for both set sizes, because in both conditions people would occasionally detect the target early during search. With the larger set size, target detection times spread over a much larger time interval, reflected in the slower growth rate of the cumulative distribution. During the RSI, rate was reduced for small but not for large set sizes. This can be explained as follows: After display offset, people might continue visual search for a few 100 ms from sensory memory, but most of the distribution's growth during the RSI probably reflects post-search processes such as response selection and motor execution, which are unaffected by set size. Therefore, the rates for large and small set sizes differed less during the RSI than before. Time pressure shortened the delay until the curves offset from zero (*a*), but slightly reduced their growth rate (*b*). This might reflect people's tendency to respond on less information under high pressure, resulting in some lucky guesses at short times, but slower accrual of accuracy over time. For small set sizes, the reduction of rate after the deadline was less pronounced under high time pressure, probably because the earlier deadline led to a larger proportion of trials for which visual search continued very briefly after display offset. This effect was not observed with the large set size because there was no reduction of rate to begin with.

To conclude, our manipulation of time pressure was effective by interrupting processing a few 100 ms after the deadline, resulting in more trials responded to incorrectly or not at all. If anything, the effect of time pressure was more severe for the larger set size, reducing overall accuracy from 91.7% correct to 68.0%, compared to a reduction from 95.1% to 81.3% for the small set size. Time pressure did not lead to a noticeable increase in information processing efficiency: The earlier offset of cumulative accuracy was compensated by a slower growth rate; this pattern reflects again the shift in time-accuracy criterion, resulting in faster responses but not a faster accrual of accuracy over time.

References

Altmann, E. M. (2009). Evidence for temporal decay in short-term episodic memory. *Trends in Cognitive Sciences, 13*, 279.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limits on retrieval. *Cognitive Psychology, 30*, 221-256.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 90-195). New York: Academic Press.

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

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

Bakeman, R., & McArthur, D. (1996). Picturing repeated measures: Comments on Loftus, Morrison, and others. *Behavioral Research Methods, Instruments, & Computers, 28,* 584-589.

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

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 33*, 570-585.

Barrouillet, P., & Camos, V. (2009). Interference: unique source of forgetting in working memory? *Trends in Cognitive Sciences, 13*, 145-146.

Barrouillet, P., De Paepe, A., & Langerock, N. (2012). Time causes forgetting from working memory. *Psychonomic Bulletin & Review, 19*, 87-92.

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.

Barrouillet, P., Portrat, S., Vergauwe, E., Diependaele, K., & Camos, V. (2011). Further evidence for temporal decay in working memory: Reply to Lewandowsky and Oberauer (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1302-1317.

Bates, D. M., Maechler, M., & Boker, B. (2011). Linear mixed-effects models using S4 classes: R Package. Retrieved from http://cran.r-project.org/web/packages/lme4/index.html

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 317-333.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

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

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.

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

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Pychology, 55*, 429-432.

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (Eds.). (2007). *Variation in working memory*. New York: Oxford University Press.

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 & Review, 12*, 769-786.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Cowan, N., & AuBuchon, A. M. (2008). Short-term memory loss over time without retroactive stimulus interference. *Psychonomic Bulletin & Review, 15*, 230-235.

Cowan, N., Elliott, E. M., Saults, J. S., Nugent, L. D., Bomb, P., & Hismjatullina, A. (2006). Rethinking speed theories of cognitive development. *Psychological Science, 17*, 67-73.

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

Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: a source activation account. *Cognitive Science, 25*, 315-353.

Evans, J. S. B. T., Handley, S., & Bacon, A. M. (2009). Reasoning under time pressure: A study of causal conditional inference. *Experimental Psychology, 56*, 77-83.

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

Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of Experimental Psychology: Learning, Memory & Cognition, 29*, 838-849.

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science, 19*, 329-335.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods, 39*, 175-191.

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

He, J., & McCarley, J. S. (2010). Executive working memory load does not compromise perceptual processing during visual search: Evidence from additive factors analysis. *Attention, Perception, & Psychophysics, 72*, 308-316.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology, 49A*, 80-115.

Hitch, G. J., Towse, J. N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General, 130*, 184-198.

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2011). How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 688-705.

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science, 6*(365-369).

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

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*, 227-229.

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 (2004). *Psychological Bulletin, 131*, 66-71.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of working memory. Mechanisms of active maintenance and executive control* (pp. 183-223). Cambridge: Cambridge University Press.

Kirkham, A. J., Breeze, J. M., & Marí-Beffa, P. (2012). The impact of verbal instructions on goal-directed behaviour. *Acta Psychologica, 139*, 212-219.

Kray, J., Eber, J., & Karbach, J. (2008). Verbal self-instructions in task switching: a compensatory tool for action-control deficits in childhood and old age? . *Developmental Science, 11*, 223-236.

Lawrence, B. M., Myerson, J., & Abrams, R. A. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin & Review, 11*, 488-494.

Lewandowsky, S. (1993). The rewards and hazards of computer simulations. *Psychological Science, 4*, 236-243.

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

Lewandowsky, S., & Farrell, S. (2011). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage.

Lewandowsky, S., Geiger, S. M., Morrell, D. B., & Oberauer, K. (2010). Turning simple span Into complex span: time for decay or interference from distractors? *Journal of Experimental Psychology: Learning Memory and Cognition, 36*(4), 958-978.

Lewandowsky, S., Geiger, S. M., & Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *Journal of Memory and Language, 59*(2), 200-222.

Lewandowsky, S., & Oberauer, K. (2008). The word-length effect provides no evidence for decay in short-term memory. *Psychonomic Bulletin & Review, 15*(5), 875-888.

Lewandowsky, S., & Oberauer, K. (2009). No evidence for temporal decay in working memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(6), 1545-1551.

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

Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009b). Response to Altmann: Adaptive forgetting by decay or removal of STM contents? *Trends in Cognitive Sciences, 13*, 280-281.

Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009c). Response to Barrouillet and Camos: Interference or decay in working memory? *Trends in Cognitive Sciences, 13*, 146-147.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*, 295-327.

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

Morey, C. C., Cowan, N., Morey, R. D., & Rouder, J. N. (2011). Flexible attention allocation to visual and auditory working memory tasks: manipulating reward induces a trade-off. *Attention, Perception, & Psychophysics, 73*, 458-472.

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

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*, 369-385.

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language, 49*(4), 469-483.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*(4), 601-626.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*(3), 544-576.

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: a computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic Bulletin & Review, 18*(1), 10-45.

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (in press). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review.*

Oriet, C., & Jolicoeur, P. (2008). Differential central resource demands of memory scanning and visual search: The role of consistent and varied mapping. *Visual Cognition, 16*, 514-551.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105*, 761-781.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220-244.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Phaf, R. H., & Wolters, G. (1993). Attentional shifts in maintenance rehearsal. *American Journal of Psychology, 106*, 353-382.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effect Models in S and S-Plus*. Berlin: Springer.

Pinheiro, J. C., Bates, D. M., DebRoy, S., & Sarkar, D. (2005). nlme: Linear and nonlinear mixed effects models. R package (Version 3.1-73).

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.

Portrat, S., Camos, V., & Barrouillet, P. (2009). Working memory in children: a time-constrained functioning similar to adults. *Journal of Experimental Child Psychology, 102*, 368-374.

R-Development-Core-Team. (2011). R: A language and environment for statistical computing Retrieved from URL: http://www.R-project.org

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

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

Ricker, T. J., & Cowan, N. (2010). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1355-1368.

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.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 419-425.

Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology, 3*, 334-349.

Svenson, O., & Edland, A. (1987). Change of preferences under time pressure: choices and judgments. *Scandinavian Journal of Psychology, 28*, 322-330.

Tubau, E., Hommel, B., & López-Moliner, J. (2007). Modes of executive control in sequence learning: From stimulus-based to plan-based control. *Journal of Experimental Psychology: General, 136*, 43-63.

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

Verbruggen, F., & Logan, G. D. (2008). Long-term aftereffects of response inhibition: Memory retrieval, task goals, and cognitive control. [Article]. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1229-1235. doi: 10.1037/0096-1523.34.5.1229

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*.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science, 21*, 384-390.

White, K. G. (2012). Dissociation of short-term forgetting from the passage of time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 255-259.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9*, 33-39.

Zelinsky, G. J., & Murphy, G. L. (2000). Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science, 11*, 125-131.

Table 1. Effects of Search Set Size at Different Levels of SOA, Experiments 1a and 1b

	SOA = 0	SOA = 0.2	SOA = 0.5	SOA = 1.0	SOA = 2.0
Experiment 1					
Set-Size Effect	493 [410, 578]	569 [491, 647]	605 [530, 680]	620 [570, 669]	659 [601, 716]
t	12.2	15.4	16.9	26.2	23.7
Proportion Reduction	.25	.14	.08	.06	--
Experiment 2					
Set-Size Effect	161 [103, 220]	160 [120, 200]	197 [137, 255]	212 [180, 245]	194 [151, 236]
t	5.8	8.4	7.0	13.8	9.5
Proportion Reduction	.17	.18	-.02	-.09	--

Note: Set-Size Effect is the RT (task 2) difference in ms between the larger and the smaller set size, with 95% confidence intervals in brackets; t refers to the t statistic (all $p < .001$), and Proportion Reduction is the proportional reduction of the set-size effect at a given level of SOA relative to the longest SOA.

Table 2. Results of ANOVAs for Response Times in Distractor-Task Trials, Experiments 3 to 5

	Experiment 3			Experiment 4			Experiment 5		
	F	p	η^2_p	F	p	η^2_p	F	p	η^2_p
Analysis of No-Switch Trials									
Before vs. During Memory	63.7	<.001	.72	11.1	.003	.33	0.2	.65	.01
Delay	991.0	<.001	.98	547.5	<.001	.96	27.8	<.001	.61
Free Time	35.1	<.001	.58	3.4	.08	.13	3.9	.06	.18
Before vs. During X Delay	14.2	<.001	.36	1.7	.20	.07	2.2	.09	.16
Before vs. During X Free Time	5.8	.02	.19	5.2	.03	.18	0.04	.83	.00
Delay X Free Time	6.9	.01	.22	1.9	.18	.08	0.9	.36	.05
3-way	0.0	.91	.00	1.8	.19	.07	1.2	.29	.06
Analysis of Trials During Memory									
Noswitch vs. Switch	32.0	<.001	.56	40.4	<.001	.64	20.8	<.001	.54
Delay	665.1	<.001	.96	272.7	<.001	.92	21.9	<.001	.55
Free Time	45.7	<.001	.65	3.3	.08	.13	4.3	.05	.19
Noswitch vs. Switch X Delay Duration	2.2	.15	.08	30.0	<.001	.57	0.0	.94	.00
Noswitch vs. Switch X Free Time	2.8	.11	.10	0.5	.51	.02	0.0	.91	.00
Delay X Free Time	1.7	.21	.06	1.2	.28	.05	3.2	.09	.15
3-way	0.5	.48	.02	1.9	.18	.08	3.6	.07	.17

Notes: Delay refers to the search set-size manipulation in Experiments 3 and 4, and to the judgment-difficulty manipulation in Experiment 5; η^2_p is partial η^2.

Table 3. Cumulative Distractor-Filled Retention Intervals in the Memory Experiments

Experiment	Short Delay	Long Delay	Percent Increase
3	15.1 (2.2)	30.2 (4.1)	101 (17)
4	23.4 (4.0)	42.5 (8.4)	82 (13)
5	17.1 (3.2)	25.7 (9.5)	50 (48)
6	24.6 (4.6)	31.3 (7.0)	27 (12)
7	23.4 (3.3)	31.1 (4.8)	32 (8)
10, low time pressure	18.2 (1.8)	33.5 (5.3)	83 (20)
10, high time pressure	14.9 (2.4)	26.6 (4.6)	80 (24)

Note: Table entries are cumulative response times to visual-search tasks, starting after presentation of the first memory item, in seconds, with standard deviations in parentheses. Short and long delays refer to conditions differing in search set size in Experiments 3, 4, and 10, and to conditions differing in distractor-task difficulty in Experiments 5-7.

Table 4: Calculation of Cognitive Load for Experiment 4

Condition	RT (2)	RT (6)	Slope	Non-search	Attn. capture	Total time	CL
Short delay (set size 2), long free time	.96		.198	.76	.528	1.76	.30
Long delay (set size 6), long free time		1.74	.198	.76	1.12	2.55	.44
Short delay (set size 2), short free time	.99		.202	.79	.545	1.19	.46
Long delay (set size 6), short free time		1.79	.202	.79	1.15	2.00	.57

Note: All times are given in seconds, estimated per search trial. RT (2) and RT (6) are response times for search set sizes 2 and 6, respectively. Slope = [RT (6) – RT (2)]/4. Non-search time = RT (2) – Slope. Attentional capture duration = 0.5 x Non-search time + 0.75 x Slope x (Setsize-1). Total time = RT + free time. Cognitive load (CL) = Attentional capture duration / Total time. All variables are calculated for each individual and then averaged.

Table 5: Cognitive Load in Experiments 3 to 5

	Short delay, long free time	Short delay, short free time	Long delay, long free time	Long delay, short free time
Exp. 3	.21 (.18)	.38 (.31)	.38 (.33)	.54 (.47)
Exp. 4	.30 (.26)	.46 (.39)	.44 (.39)	.57 (.48)
Exp. 5	.27 (.17)	.46 (.29)	.40 (.35)	.60 (.54)

Note: Entries in parentheses are conservative estimates of cognitive load on the assumption of strategic postponement of distractor task responses for refreshing.

Table 6: Memory Accuracy in Experiments 3 to 5: Experimental Data and Simulations with TBRS*

Data Source	Short delay, long free time	Short delay, short free time	Long delay, long free time	Long delay, short free time	Effect of delay (95% CI)
Experiment 3					
Exp.	.93 (.05)	.91 (.09)	.91 (.10)	.93 (.09)	-.001 [-.021, .020]
Sim.	.78	.71	.66	.56	.14
Experiment 4					
Exp.	.67 (.28)	.68 (.21)	.69 (.23)	.67 (.21)	-.005 [-.041, .032]
Sim.	.74	.64	.35	.37	.34
Experiment 5					
Exp.	.89 (.08)	.90 (.08)	.91 (.08)	.90 (.01)	-.006 [-.039, .011]
Sim.	.77	.72	.66	.49	.17

Note: Exp. refers to experimental data, and Sim. to results of the TBRS* simulation. Entries in parentheses are standard deviations. In the last column, 95% confidence intervals for the observed delay effect are given in brackets. Parameters for the TBRS* simulation were: Position overlap = 0.3; criterion for encoding $\tau_E = .95$, processing rate $R = 6$; SD of processing rate $s = 1$; decay rate $D = 0.5$; refreshing duration $T_r = 80$ ms; threshold for retrieval $\theta = 0.1$; noise $\sigma = 0.02$.

Table 7: Results of Experiments 6 and 7

	Easy, long free time	Easy, short free time	Hard, long free time	Hard, short free time	Effect of decision difficulty (and delay)
Distractor-Task Proportion Correct					
Experiment 6	.99 (.02)	.99 (.01)	.79 (.12)	.79 (.14)	.20
Experiment 7	.98 (.02)	.98 (.01)	.84 (.07)	.83 (.08)	.15
Distractor-Task Response Times (s)					
Experiment 6	1.00 (.19)	1.06 (.21)	1.29 (.31)	1.34 (.30)	0.28
Experiment 7	0.96 (0.14)	0.99 (0.15)	1.25 (0.19)	1.34 (0.22)	0.32
Memory Proportion Correct					
Experiment 6	.92 (.05)	.93 (.07)	.90 (.10)	.91 (.09)	.018 [-.007, .044]
Experiment 7	.87 (.15)	.88 (.13)	.88 (.17)	.84 (.16)	.013 [-.026, .051]

Note: Entries in parentheses in the first four data columns are standard deviations; in the final column the 95% confidence interval for the effect of difficulty on memory proportion correct is given in brackets.

Table 8. Results of ANOVAs for Response Times in Distractor-Task Trials, Experiments 8 and 9

	Experiment 8			Experiment 9		
	F	p	η^2_p	F	p	η^2_p
Analysis of No-Switch Trials						
Before vs. During Memory	43.3	<.001	.71	38.2	<.001	.34
Time Pressure	53.0	<.001	.75	121.2	<.001	.62
Free Time	1.5	.24	.08	1.1	.301	.01
Before vs. During X Time Pressure	10.9	.004	.38	35.5	<.001	.32
Before vs. During X Free Time	9.0	.008	.33	4.2	.044	.05
Time Pressure X Free Time	28.3	<.001	.59	0.3	.58	.00
3-way	0.2	.63	.01	9.6	.003	.12
Analysis of Trials During Memory						
Noswitch vs. Switch	10.1	.005	.36	28.2	<.001	.28
Time Pressure	56.9	<.001	.76	179.2	<.001	.71
Free Time	6.5	.020	.27	7.2	.009	.09
Noswitch vs. Switch X Time Pressure	4.3	.051	.20	17.4	<.001	.19
Noswitch vs. Switch X Free Time	2.0	.18	.10	53.0	<.001	.42
Time Pressure X Free Time	23.0	<.001	.56	6.5	.01	.08
3-way	2.0	.10	.14	0.4	.52	.01

Notes: Free Time is the time following each visual-search trial before onset of the next stimulus in Experiment 8, and the temporal gap between memory items and first visual-search display in Experiment 9; η^2_p is partial η^2.

Table 9: Memory Accuracy in Experiments 8 and 9

	Low Pressure, long free time	Low pressure, short free time	High pressure, long free time	High pressure, short free time	Time pressure effect
Exp. 8	.89 (.09)	.84 (.09)	.85 (.12)	.76 (.14)	.06 [.024, .100]
Exp. 9	.76 (.18)	.74 (.19)	.68 (.20)	.68 (.19)	.07 [.042, .093]

Note: Free Time refers to the time following each visual-search trial before onset of the next stimulus in Experiment 8, and the temporal gap between memory items and first visual-search display in Experiment 9. Entries in parentheses are standard deviations. In the last column, 95% confidence intervals for the time-pressure effect are given in brackets.

Table 10: Results of Experiment 10

	Short delay, low pressure	Short delay, high pressure	Long delay, low pressure	Long delay, high pressure	Effect of delay
Distractor RT	.77 (.08)	.70 (.08)	1.43 (.22)	1.18 (.21)	0.57
Cognitive load	.34 (.02)	.35 (.03)	.48 (.03)	.47 (.05)	.13
Proportion correct memory	.85 (.11)	.80 (.14)	.83 (.14)	.80 (.13)	.013 [-.008, .034]

Note: Distractor RT and cognitive load were computed from the second sub-sample for which responses were collected during the RSI. For visual-search trials without response, RT was set to the maximum available time (i.e., deadline + RSI). Proportion correct memory is given for all participants. Entries in parentheses in the first four columns are standard deviations; the last column includes the 95% confidence intervals for its effect on memory (in brackets).

Table 11: Parameter Estimates for Fixed Effects from Mixed-Effect Model of Recall in Experiment 10

	Full Model (BIC = 5523)		Final Model (BIC = 5508)	
Predictor	Estimate	SE	Estimate	SE
Search set size	-.54	.16		
Time pressure	1.17	.29	.99	.28
Number of processing bursts with time-outs	-.19	.03	-.21	.02
Number of additional time-outs	-.03	.01		
Set size x time pressure	-.52	.47		

Table A1: Parameter Estimates for Fixed Effects of Mixed-Effect Model for Cumulative Distributions of Correct Responses to Visual Search in Experiment 10.

Parameter	Set size 2, low pressure	Set size 2, high pressure	Set size 6, low pressure	Set size 6, high pressure
a	.56	.49	.72	.58
b	4.09	3.96	1.17	1.04
br	1.93	2.80	1.24	1.00
c	.999	.999	.999	.999

Note: Estimated curve parameters were calculated from the intercept and slope fixed effects of the mixed-effect model.

Figure Captions

Figure 1. Hypothetical schedules of processing steps of task 1 and task 2 in a PRP paradigm, assuming that the visual-search component of task 2 must wait until the central component of task 1 has been completed (top), or assuming that the visual-search component can run in parallel with the central component of task 1 (bottom). Each scenario depicts the processing stages of task 1, together with four conditions of task 2. The first two conditions are small vs. large set size at SOA = 0, and the final two conditions are small vs. large set size at a very long SOA. Stages are S = sensory, RS = response selection, M = motor, VS = visual search (with set size). Stages assumed to require the central attentional bottleneck are filled black.

Figure 2. Example of a visual-search display with set size six.

Figure 3. Mean reaction times for task 2 in the PRP experiments (Experiment 1 in the top panel and Experiment 2 in the bottom panel), by visual search set size and SOA. Error bars represent 95% confidence intervals for within-subject comparisons (Bakeman & McArthur, 1996).

Figure 4. Variant of the scenario depicted in the bottom panel of Figure 1: The search component for the large set size takes longer than the waiting time at SOA = 0. As a consequence, a small set-size effect remains at SOA = 0, but at the same time, there is no waiting time left for the large set size.

Figure 5. A: Sequence of events in a complex-span trial. Bursts of four distractor trials (visual search) alternate with encoding of memory items (letters K and F), until question marks probe for recall after the last burst. B: Hypothetical sequence of processing stages within each visual-search trial, for easy (set size 2) and hard (set size 6) conditions. Onset of the search display is followed by four processing stages, sensory processing (S), visual search (VS), response selection (RS), and motor execution (M),

after which the screen is blank for a fixed RSI (free time). Black sections represent periods during which the attentional bottleneck is engaged (for at least 75% of the time during VS); white sections indicate periods during which attention is free for presumed memory restoration processes.

Figure 6. Mean reaction times for visual search in the context of the complex span task (Experiments 3 and 4), for small and large search-set sizes, separately for (no-switch) trials before presentation of the first memory item (“before”), no-switch trials during presentation of memory items (“during”), and switch trials during presentation of memory items (“switch”). Error bars represent 95% confidence intervals for within-subject comparisons (Bakeman & McArthur, 1996).

Figure 7. Decomposition of the time in one visual-search trial, as used for the computation of cognitive load. The black box at the beginning represents encoding of the preceding memory item (M). The black areas are the times assumed to require central attention; the white areas are times assumed to be free for refreshing because the attentional bottleneck is not involved in those processes. The area with the padlock represents the experimenter-controlled free time added after the response (black vertical bar). Cognitive load is the grey area (attentional capture) divided by the sum of the grey and the white areas (total time). The top panel (a.) shows our standard method for computing cognitive load. Non-search time is subdivided into response selection (RS, 50%) on the one hand, sensory (Sens.) and motor execution (Mot.) times (together 50%) on the other hand. Response selection is known to involve the bottleneck unlike both sensory and motor processes, Search time (VS) is subdivided into search without central attention (25%) and search with central attention (75%), as determined by Experiments 1 and 2. The bottom panel (b.) illustrates the conservative estimate of cognitive load which involves two more estimates of time during which the attentional bottleneck is available for refreshing: Switch cost is the RT difference between the

first and the following visual-search trials in a burst. Delay is the difference between no-switch trials during memory and no-switch trials in the pre-burst. These times are taken to be available for memory refreshing on the assumption that they reflect participants' choice to withhold a response for that purpose. (The switch cost is added only once per burst; the remaining three search trials don't include it). The proportions of areas in the panels roughly correspond to the proportions of times measured in the first search trial of each burst in the large set-size (long delay), short free-time condition of Experiment 3.

Figure 8. Serial position curves for recall of letters in correct position by search set size and free time, Experiments 3 (top panel) and 4 (bottom). Error bars represent 95% confidence intervals for within-subject comparisons.

Figure 9. Proportion correct in Experiments 3 to 5 as a function of cognitive load. Cognitive load estimates are the ones without parentheses in Table 5.

Figure 10. Stimuli for the spatial fit-judgment task. Stimuli were centred on the screen. Distance between small rectangles varied in 6 steps from 1/15 to 1/30 of the horizontal screen extension (X). The bar appeared above the gap between the two rectangles in half the trials, and below the gap in the other trials; the distance between the bar and the gap was 1/15 of vertical screen extension (Y). In difficult trials, the horizontal bar extension differed from a bar just fitting into the gap by plus or minus 2 pixels; in the easy condition, it differed by plus or minus 20 pixels.

Figure 11. Relationship between total time, free time, attentional capture duration and time pressure in Experiment 8. In all panels, the vertical bar labelled "R" refers to the time of overt responding, the broken vertical line labelled "DL" is the deadline. Box labelled "M" represents a list item, and boxes "D1" and "D2" represent two distractors. The grey horizontal bars with the padlock symbol represent the experimenter-controlled free-time interval added after display offset (at the response

or the deadline, whichever comes first). The effective free time also includes some proportion of processing time (unfilled bar) during which attention is not captured, and estimates of cognitive load are corrected for this component. Panels (a) and (b) show two trials (one fast, one slow, respectively) with low time pressure; panels (c) and (d) show two trials (also fast and slow) with high time pressure. With high time pressure, the slower trial cannot be completed in time, and therefore a portion of attentional capture duration spills into the free-time period. On balance, therefore, increasing time pressure can be expected to leave cognitive load unaffected.

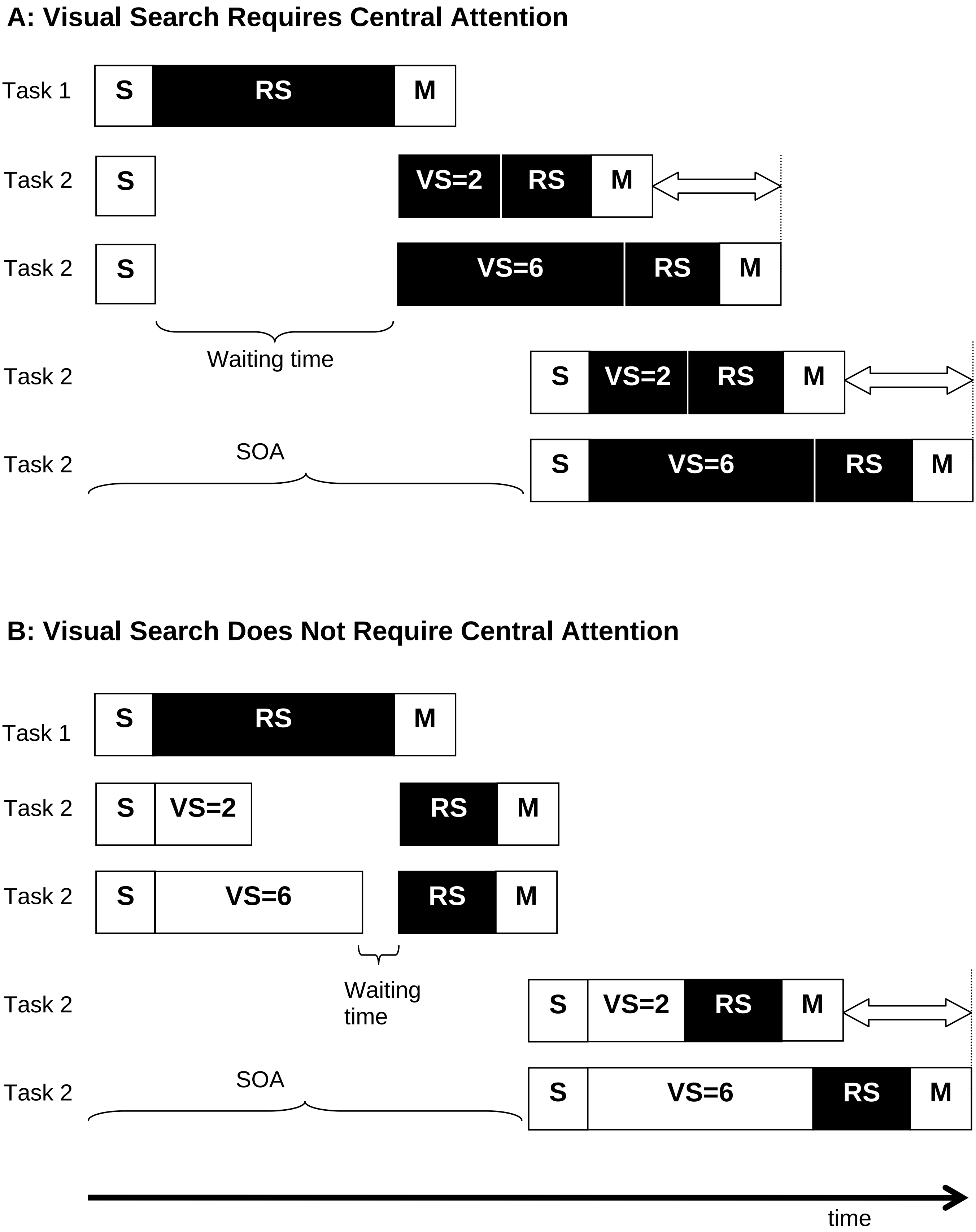
Figure 12. Response times to visual-search distractors in Experiments 8 and 9. Error bars represent 95% confidence intervals for within-subject comparisons.

Figure 13. Manipulation of time pressure and attentional-capture duration in Experiment 10. Panel (a.) has a small search set size which leads to relatively brief attentional capture and moderate cognitive load. Panel (b.) has a larger set size and hence greater cognitive load because free time is roughly comparable across panels. Time pressure is constant across both situations because the deadline (“DL”) is in the same position relative to the individually-calibrated distractor processing times. In consequence, time pressure can be dissociated from cognitive load. In panels (c.) and (d.) time pressure is increased for small and large search set sizes, respectively. The increase in time pressure does not moderate the effect of set size, and hence does not alter the differences in cognitive load between (c.) and (d.).

Figure 14: Memory performance in Experiment 10 as a function of cognitive load and time pressure. Error bars represent 95% confidence intervals for within-subject comparisons.

Figure A1. Cumulative distributions of correct responses to visual-search distractors in Experiment 10 (second sample only). Points represent data; lines reflect the fitted model. Vertical lines are the mean deadlines for high time pressure (left line) and low time pressure (right line).

Figure 1



A: Visual Search Requires Central Attention

B: Visual Search Does Not Require Central Attention

Figure 2

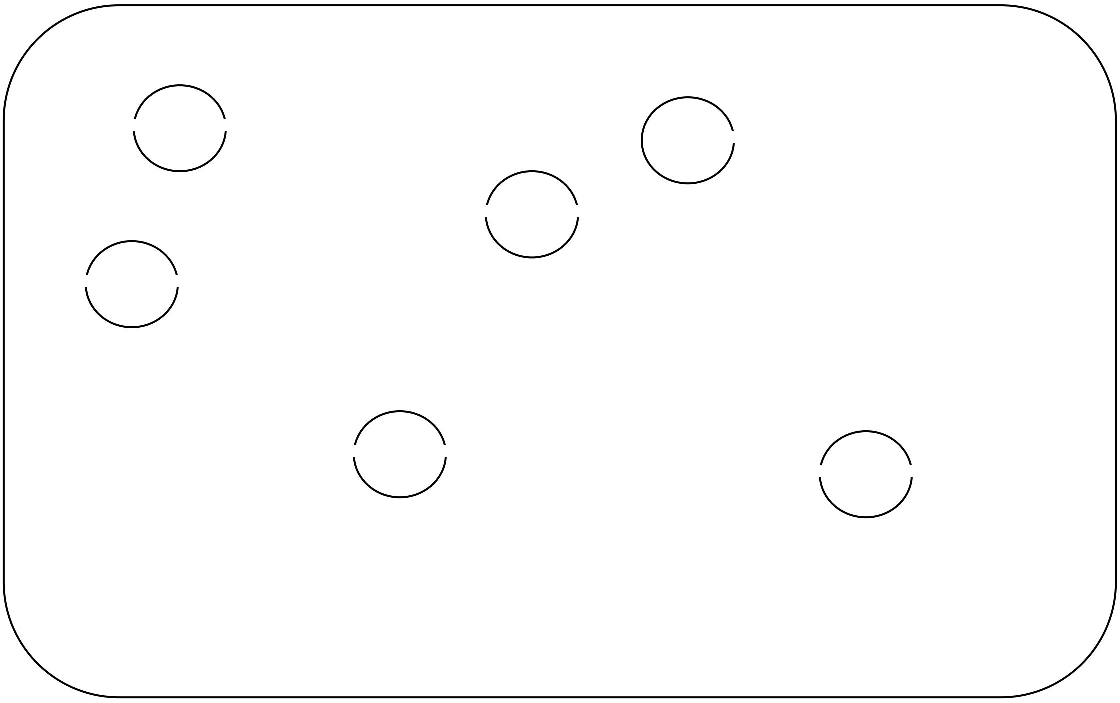


Figure 3

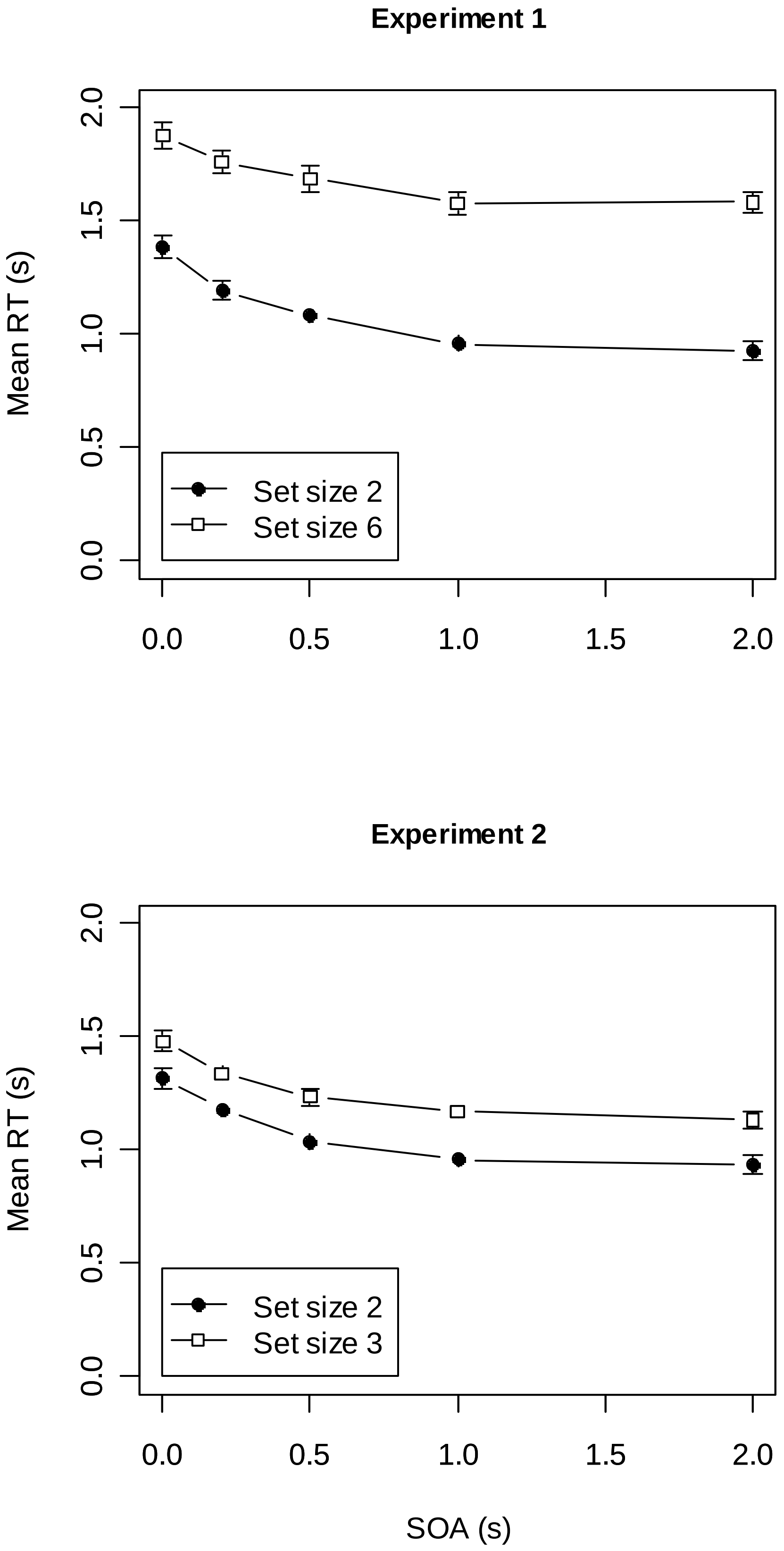

Experiment 1
Mean RT (s)
0.0
0.5
1.0
1.5
2.0
Set size 2
Set size 6
0.0
0.5
1.0
1.5
2.0
Experiment 2
Mean RT (s)
0.0
0.5
1.0
1.5
2.0
Set size 2
Set size 3
0.0
0.5
1.0
1.5
2.0
SOA (s)

Figure 4

B’: Visual Search Does Not Require Central Attention; Search is not completely absorbed by waiting time

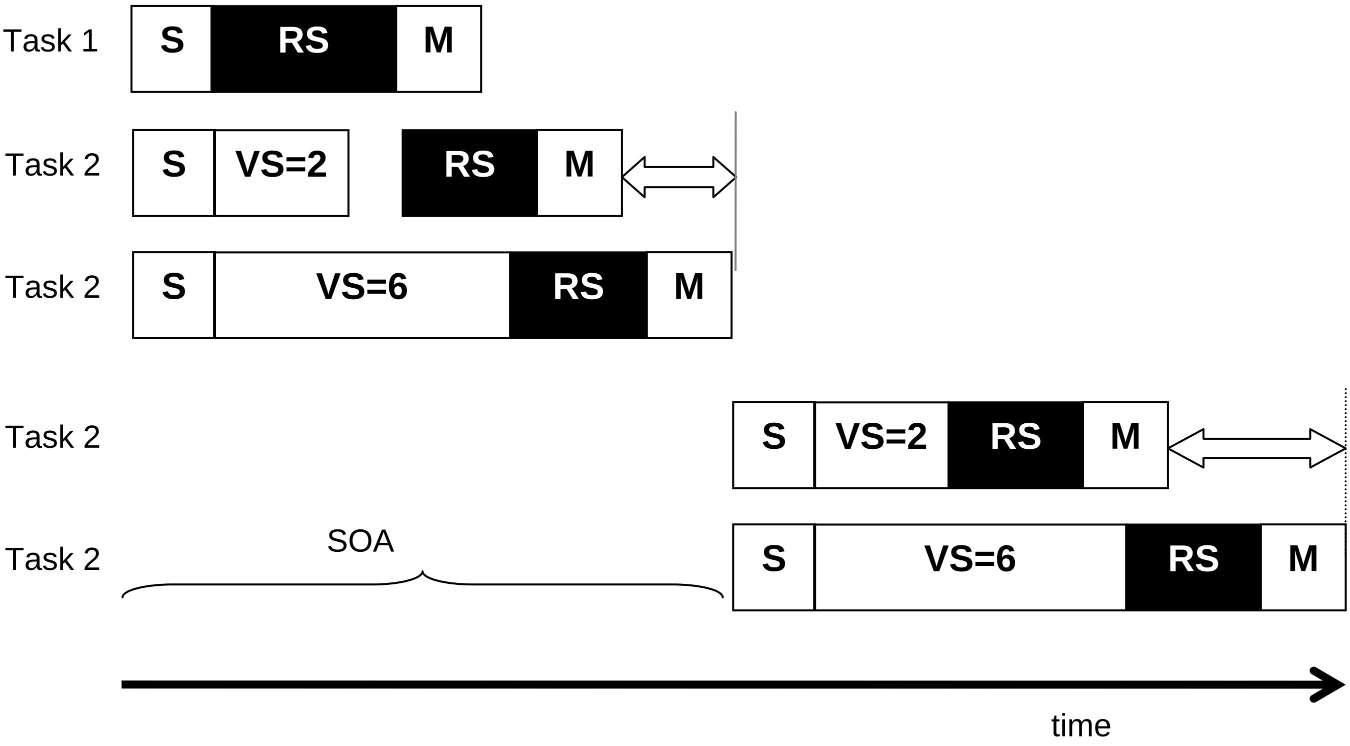


Figure 5

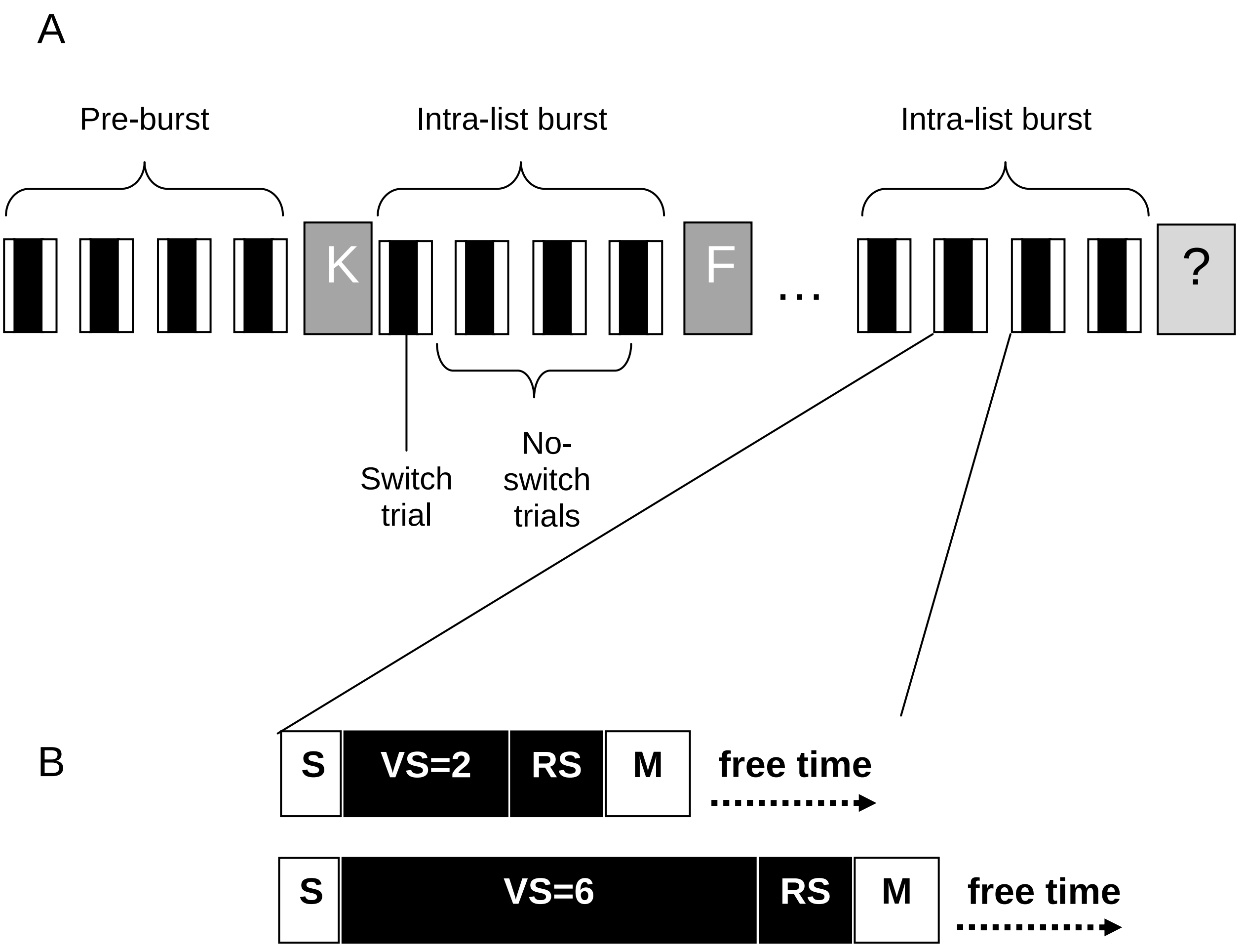

A
Pre-burst
Intra-list burst
Intra-list burst
K
F
…
?
Switch trial
No-switch trials
B
S
VS=2
RS
M
free time
S
VS=6
RS
M
free time

Figure 6

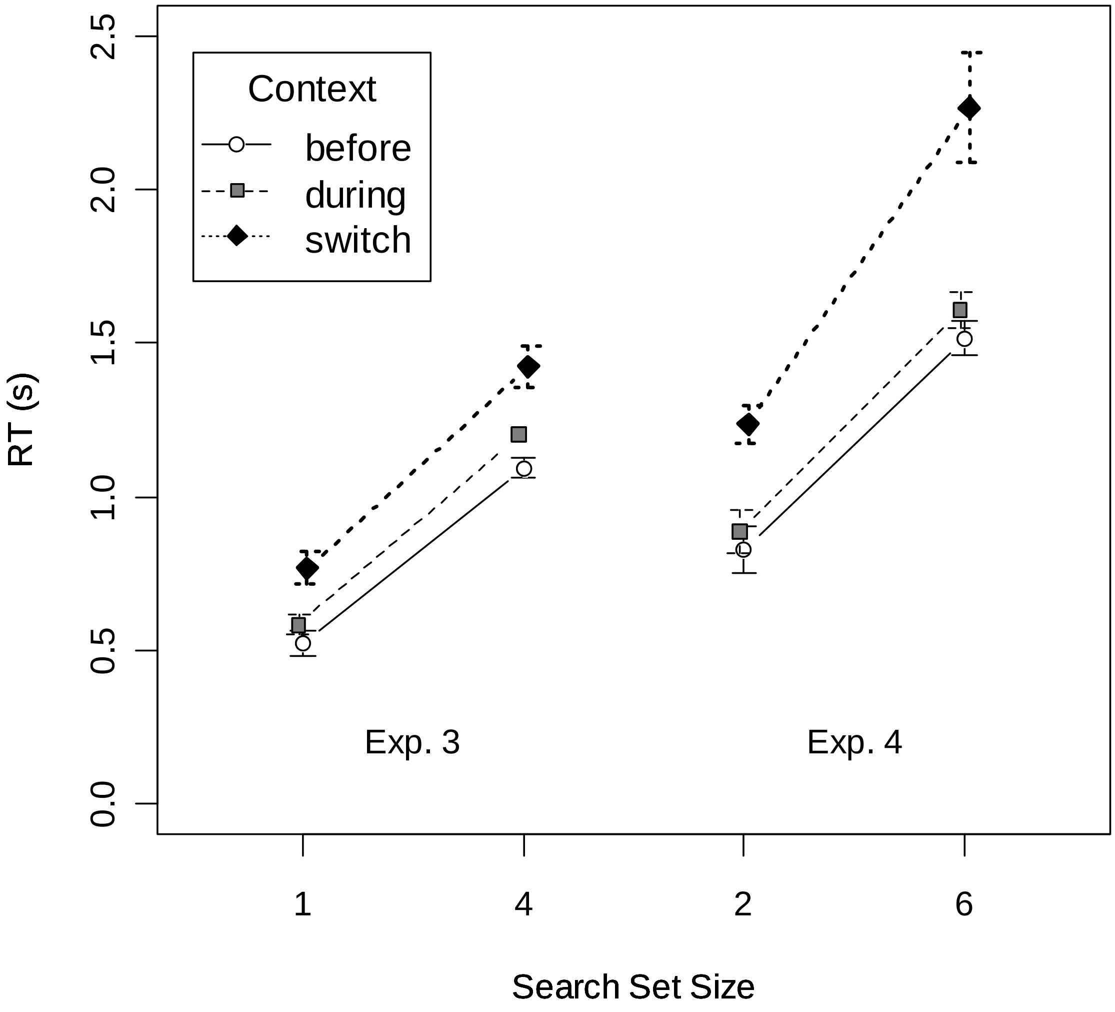

Context
before
during
switch
RT (s)
0.0
0.5
1.0
1.5
2.0
2.5
Exp. 3
Exp. 4
1
4
2
6
Search Set Size

Figure 7.

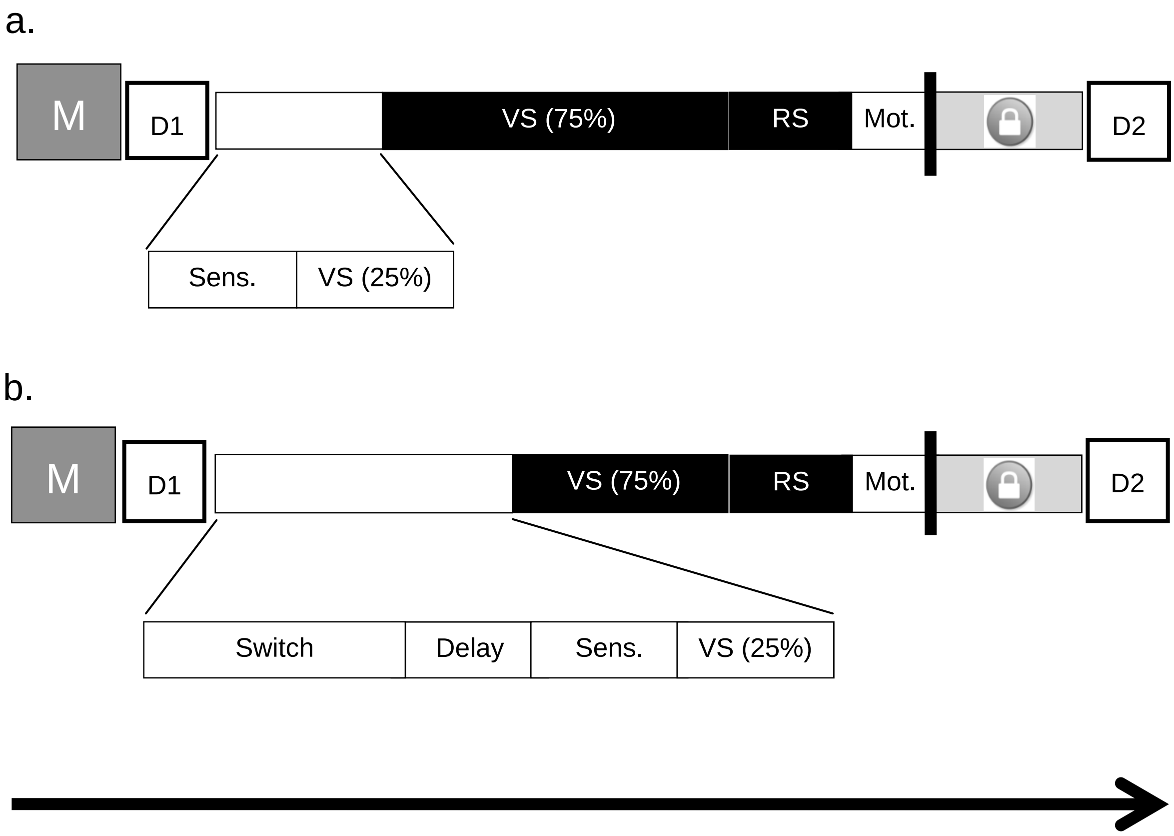

a.
M
D1
VS (75%)
RS
Mot.
D2
Sens.
VS (25%)
b.
M
D1
VS (75%)
RS
Mot.
D2
Switch
Delay
Sens.
VS (25%)

Figure 8

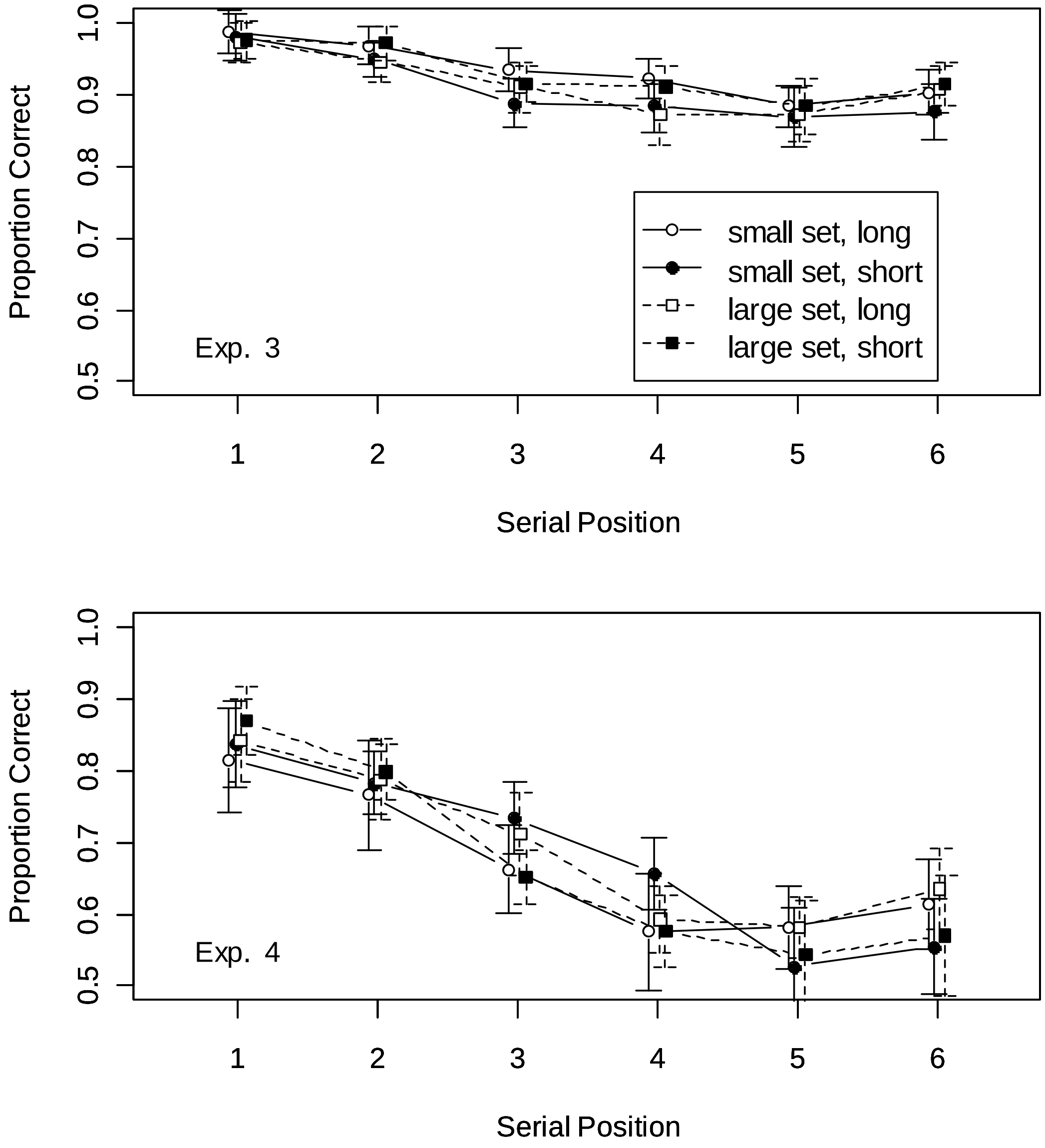

Proportion Correct
0.5
0.6
0.7
0.8
0.9
1.0
Exp. 3
small set, long
small set, short
large set, long
large set, short
1
2
3
4
5
6
Serial Position
Proportion Correct
0.5
0.6
0.7
0.8
0.9
1.0
Exp. 4
1
2
3
4
5
6
Serial Position

Figure 9

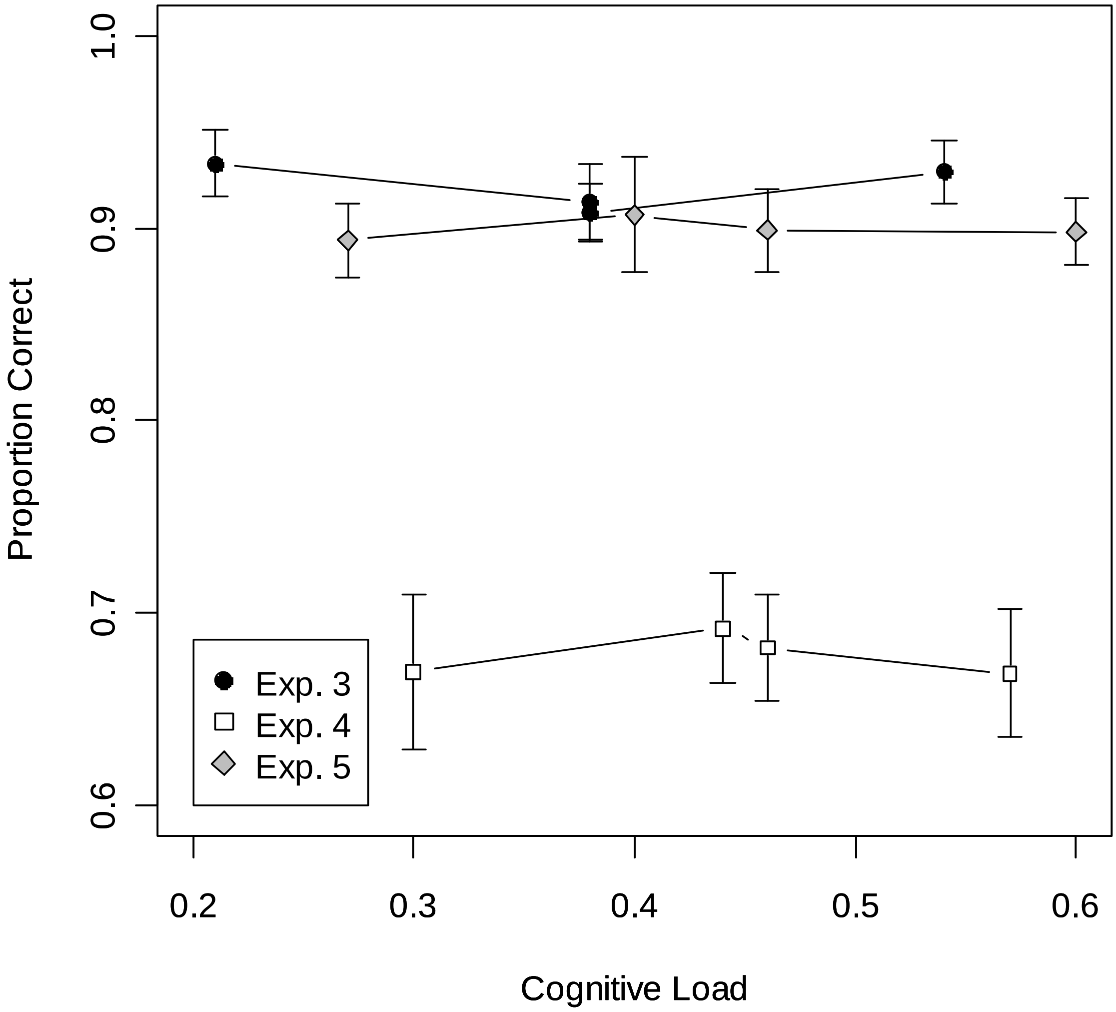
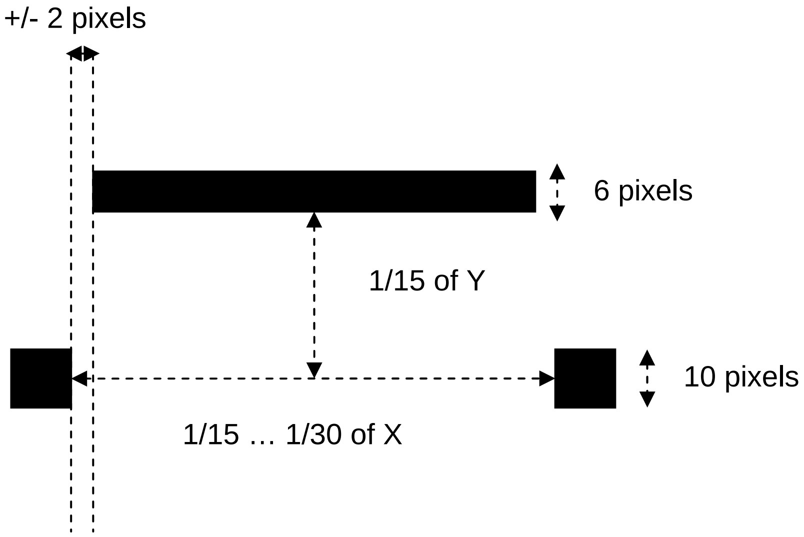

1.0
0.9
0.8
0.7
0.6
Proportion Correct
0.2
0.3
0.4
0.5
0.6
Cognitive Load
Exp. 3
Exp. 4
Exp. 5

Figure 10

Difficult



Easy

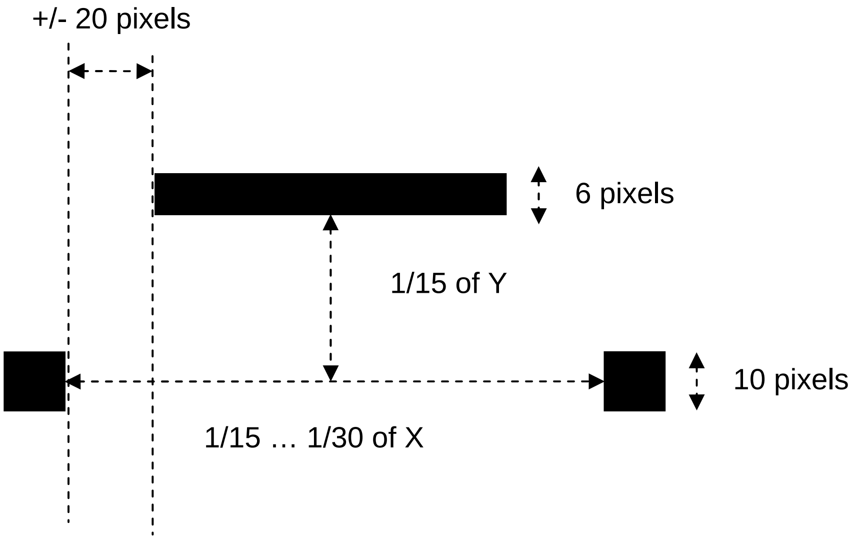


Figure 11

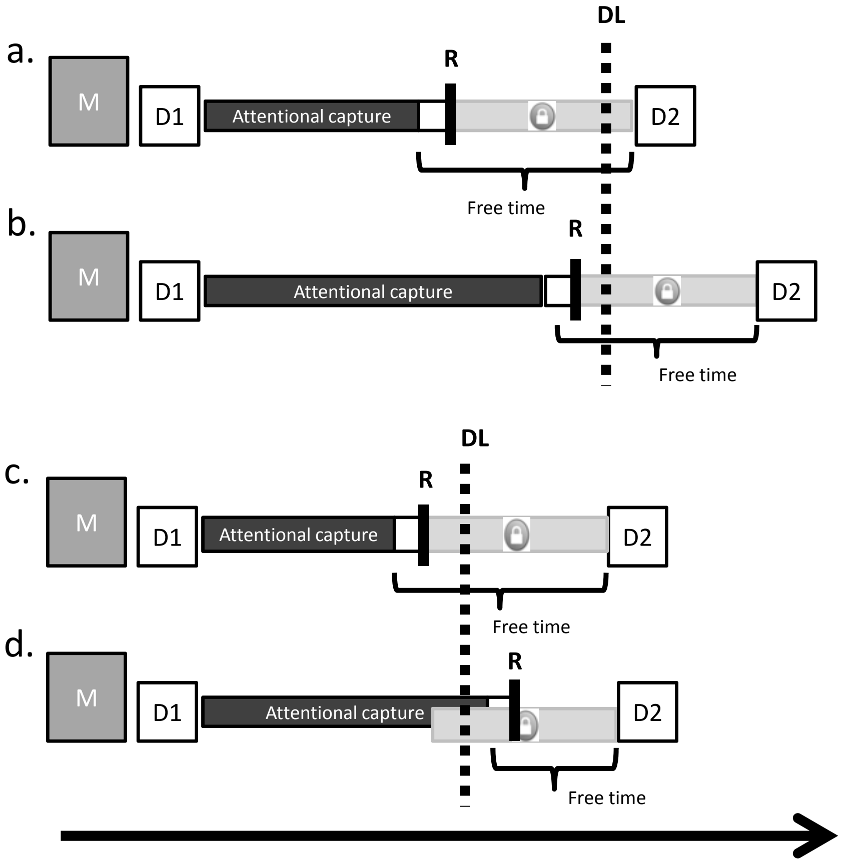

DL
a.
R
M
D1
Attentional capture
D2
Free time
b.
R
M
D1
Attentional capture
D2
Free time
DL
c.
R
M
D1
Attentional capture
D2
Free time
d.
R
M
D1
Attentional capture
D2
Free time

Figure 12

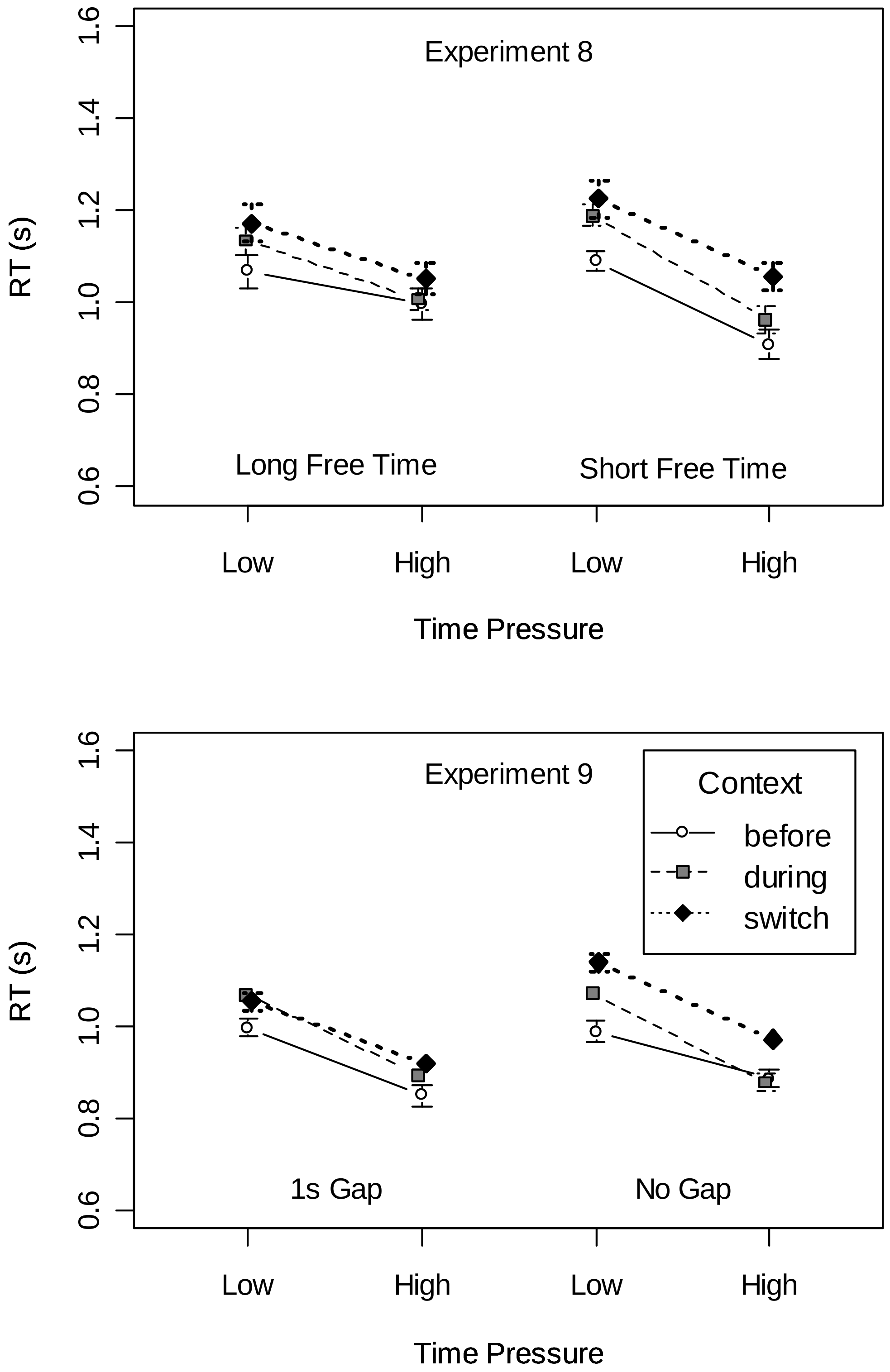

Experiment 8
RT (s)
1.6
1.4
1.2
1.0
0.8
0.6
Long Free Time
Short Free Time
Low
High
Low
High
Time Pressure
Experiment 9
Context
before
during
switch
RT (s)
1.6
1.4
1.2
1.0
0.8
0.6
1s Gap
No Gap
Low
High
Low
High
Time Pressure

Figure 13

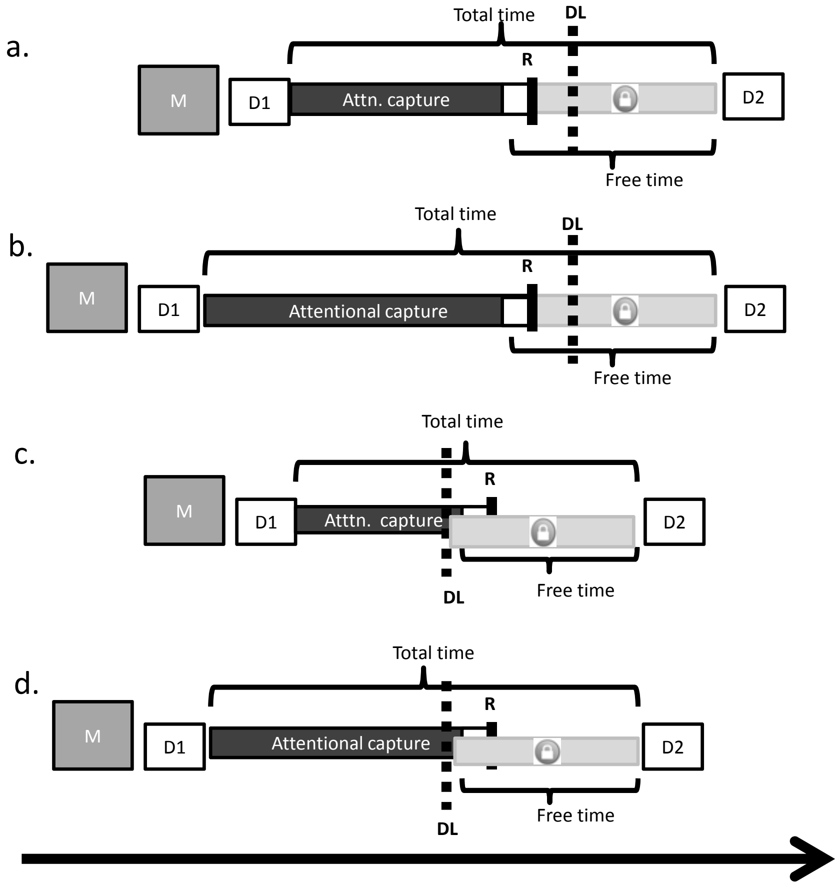

a.
Total time
DL
R
M
D1
Attn. capture
D2
Free time
b.
Total time
DL
R
M
D1
Attentional capture
D2
Free time
c.
Total time
R
M
D1
Atttn. capture
D2
DL
Free time
d.
Total time
R
M
D1
Attentional capture
D2
Free time
DL

Figure 14

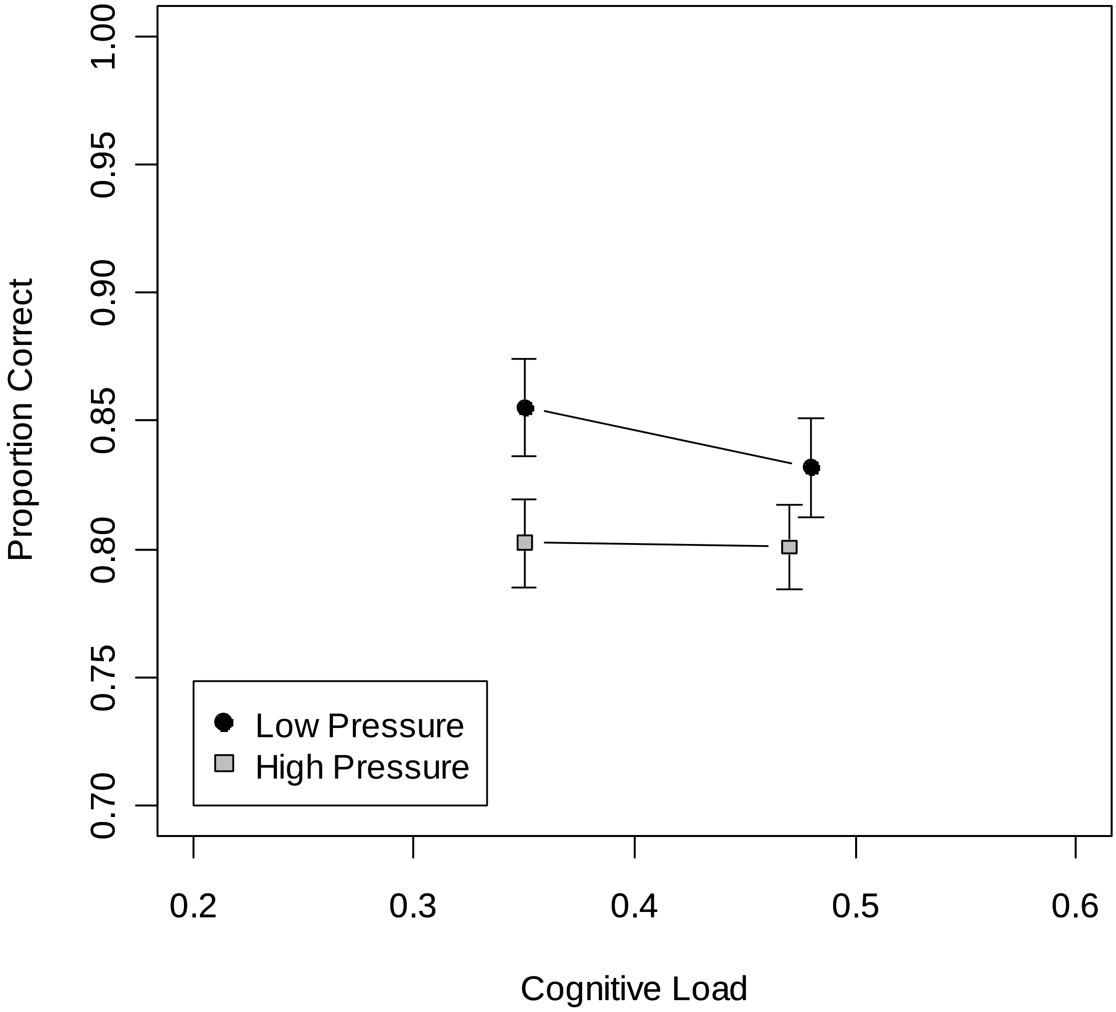
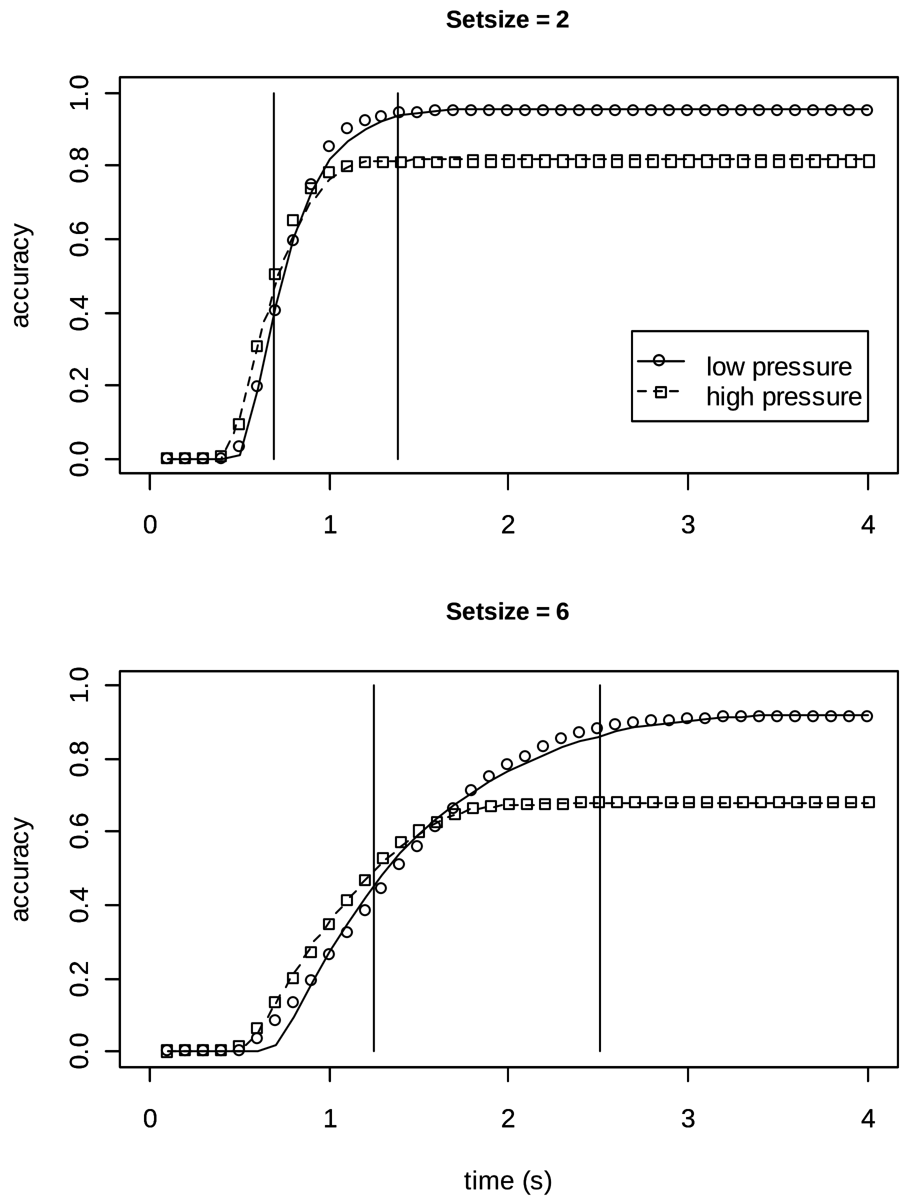

Proportion Correct
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.2
0.3
0.4
0.5
0.6
Cognitive Load
● Low Pressure
■ High Pressure

Figure A1



Setsize = 2

Setsize = 6