4

The time-based resource-sharing model of working memory

Pierre Barrouillet and Valérie Camos

Since Baddeley and Hitch's (1974) seminal proposal to distinguish between short-term and working memory, several influential models and theories have been put forward to account for the structure and functioning of a working memory thought of as the core system of cognition devoted to the maintenance and processing of relevant information. As Engle and Oransky (1999) suggested, deas about temporary storage in memory during the previous decades have moved from multi-store models and structural approaches championed by Baddeley and Hitch (1974; Baddeley and Logie 1999) to more dynamic or process-oriented models of working memory (Cowan 1995 1999; Engle, Cantor and Carullo 1992; Engle, Kane and Tuholski 1999; Lovett, Reder and Lebiere 1999; Schneider and Detweiller 1987). By conceiving working memory as the activated part of long-term memory rather than as a structure or a set of cognitive structures, these theories have emphasized functional aspects of cognition related to the way activation is produced, maintained or even inhibited, and to the way cognitive processes use the activated items of knowledge. The main question, which seemed to be 'How is working memory organized?' has progressively become 'How does working memory work?'. We do not claim to know how working memory works and admit that the reader will not find the final answer in this chapter. However, we hope to provide some elements to shed light on questions such as 'How is the relevant information maintained active during concurrent processing?', 'What is the nature of the resource to be shared between maintenance and processing?', 'How is this sharing achieved?', and thus 'What are the limiting factors of working memory functioning?'.

Retracing the history of the literature on short-term and working memory, Engle and Oransky (1999) noted that models have become more complex and more flexible. Whereas the first descriptions of working memory functioning mainly concerned the articulatory loop, which is closest in character to the original concept of a short-term store, further models provided thorough descriptions of the relationships between short- and long-term memory, of processes devoted to the activation and maintenance of information, as well as of the role of attention in selecting relevant information and monitoring working memory functioning. Several complex span tasks designed to evaluate working memory capacity proved to be reliable and predictive of high level cognition, and the relations between working memory capacity, controlled attention and general intelligence have been investigated extensively. However, these models might have overlooked an important factor in accounting for cognition, namely time.

Many models of working memory mentioned time as an important factor, but almost exclusively to account for forgetting from short-term memory. For example, Baddeley (1990, 2000) suggested that the traces within the articulatory or phonological loop decay within about 2 s unless they are maintained by a process of subvocal articulatory rehearsal. Cowan (1995) assumed that activation of the information that is outside the focus of attention declines with time. This hypothesis of a time-related decay of memory traces has even been the basis of a

recent account of the processes underlying performance on working memory span tasks put forward by Towse and Hitch (1995; Hitch, Towse and Hutton 2001; Towse, Hitch and Hutton 1998, 2000, 2002). Working memory span tasks are complex span tasks in which the participants have to maintain some to-be-remembered items while performing a concurrent activity. For example, in the counting span task (Case, Kurland and Goldberg 1982), the participants are asked to count dots on cards out loud, and then recall the number of dots on each card. The experimenter varies the number of cards to be counted, and consequently the number of values to be recalled after counting. The counting span is the maximum number of cards the participant is able to remember. According to Towse and Hitch's (1995) hypothesis, the difficulty of the working memory span tasks is due to the fact that the memory traces of the items to be recalled suffer from a time-related decay. Because participants have to switch their attention from processing to storage when performing working memory span tasks, their recall performance would depend on the duration of the processing which in turn determines the retention period during which memory traces fade away (see also Halford, Maybery, O'Hare *et al.* 1994). In all of these theories, time affects maintenance because traces are lost over time through decay, but this time parameter does not affect working memory functioning per se. Even in Towse *et al.*'s (1998) proposals, the switches of attention between the processing and maintenance parts of the complex span tasks are assumed to occur in a simple manner that reflects the way working memory span tasks are structured: attention would be entirely absorbed by counting and there would be no active maintenance of stored totals that competes with the execution of counting operations.

In contrast, we assume that the dual function of a working memory devoted to the active maintenance of information while concurrent processing is performed demands a subtle interplay between activities in which time plays the crucial role. The aim of our time-based resource-sharing model (Barrouillet, Bernardin and Camos 2004) is to describe the time course of working memory functioning, the constraints and limits that affect this functioning, how these limits differ from one individual to another and how they evolve with age. We assume that working memory span tasks constitute an appropriate and in fact ideal model to study working memory functioning. Their dual-task structure perfectly mimics the main constraint our cognitive system is faced with: the need to maintain active and ready for treatment transient short-lived memory items while performing sometimes complex and time-consuming activities. Thus, the time-based resource-sharing model has been developed and tested using traditional but also new working memory span tasks in which time parameters are carefully controlled.

The time-based resource-sharing model: the main proposals

Our model is based on four main proposals. First, we assume that both processing and maintenance require attention, which is a limited resource and that, as a consequence, some sharing is needed. Second, following Cowan (1995) and Towse and Hitch (1995), as soon as attention is switched away, the activation of the memory traces suffers from a time-related decay. Refreshing these decaying memory traces requires their retrieval from memory by attentional focusing. Third, as a consequence, any processing that captures attention would disrupt maintenance by preventing the refreshment of these memory traces through attentional focusing. In a first version of the theory (Barrouillet *et al.* 2004), it was assumed that among the activities that capture attention, those that occupy the retrieval process needed to refresh memory traces should have an especially detrimental effect on maintenance because there is a central bottleneck that constrains retrieval activities. This precise point will be addressed here and we shall see that working memory functioning is limited by a central bottleneck that does not only constrain retrievals but probably any central process. Fourth, because the central bottleneck allows only one central process at

a time, sharing attention is time-based. It is assumed that within tasks that require attention there are periods of time during which attention is totally captured, preventing memory traces' refreshment. Thus, attention sharing is achieved through a rapid and frequent switching between processing and maintenance that occurs during the completion of the task.

As far as the first point is concerned, the time-based resource-sharing model assumes that processing and maintenance rely on the same limited attention resource. Most of the main working memory span tasks involve complex activities as processing components (e.g., reading sentences, counting dots, arithmetic equation solving, reasoning). These complex activities often require planning multistep strategies, setting goals and subgoals, and maintaining intermediary results. All of these activities are known to require controlled attention, which is a limited resource considered by Engle as isomorphic with fluid intelligence (Engle *et al.* 1999b; Conway, Kane and Engle 2003). The complex activities used in working memory span tasks require attention not only because they involve high level cognition but also in their most elementary processing steps such as retrievals of knowledge from long-term memory. Indeed, within Anderson's Adaptive Character of Thought (ACT-R) framework, activation of declarative knowledge from long-term memory is achieved through attentional focusing on retrieval cues. Interestingly, Anderson (1993) assumes that the total amount of attention considered as source of activation is limited and probably varies from one individual to another (Lovett *et al.* 1999). Thus, the processing component of most of the working memory span tasks require attention. Moreover, several models of working memory assume that maintenance is also attention-demanding. In line with these specific models, we assume that short-term memory is that part of long-term memory attentionally activated above threshold (Anderson 1993; Cantor and Engle 1993; Cowan 1995; Engle and Oransky 1999; Lovett *et al.* 1999). Working memory would thus contain those items on which attention focuses as well as items that are outside of the focus of attention but sufficiently activated to be readily accessible (Conway and Engle 1994; Cowan 1995). As we noted above, activating items from long-term memory requires attention. Thus, both components of working memory span tasks, and more generally the two main functions of working memory that are maintenance of information and processing rely on the same limited attention resource. As a consequence, some sharing of this resource is needed.

As far as the second point is concerned, the items on which attention is focused receive activation, but this activation decays as soon as attention is switched away (Anderson 1993; Cowan 1995, 1999). As a consequence, the memory items to be recalled in working memory span tasks suffer from a time-related decay when attention is switched away from maintenance (Towse and Hitch 1995). The refreshment of these items before their complete disappearance necessitates their reactivation by attentional focusing. In other words, maintaining items in short-term memory in view of their further recall requires the individual to frequently switch attention away from processing to prevent forgetting.

Thus, and this is our third point, memory decay and, consequently, recall performance would be a function of the time during which the concurrent processing captures attention and thus impedes refreshment. As suggested by Barrouillet and Camos (2001), any processing component that captures some part of the available attention capacity should have this detrimental effect. Whether the detrimental effect of the processing component on recall is due to the necessity to share between processing and maintenance a specific central process such as memory retrieval, or just to the fact that the processing component captures attention and thus impedes refreshment remains an open question. Indeed, the effect of processing on maintenance could be accounted for either by the hypothesis of a bottleneck for retrieval or by the hypothesis of a capture of attention. On the one hand, many studies suggest that when two concurrent tasks rely on the same central process, severe interference can be observed. For example, Rohrer, Pashler and

Etchegaray (1998) demonstrated that two memory retrievals cannot be performed simulta-neously, suggesting that there is a central bottleneck for retrieval (Pashler 1998). To account for this fact, the authors suggested the metaphor of a spotlight of retrieval that would be directed at only one category at a time in such a way that multiple categories can simultaneously remain acti-vated for a short while, but only one category can be lit. Thus, in the first version of our model, we assumed that the detrimental effect of processing on maintenance would be particularly pro-nounced when the processing component involves memory retrievals, which is the case in the main verbal working memory span tasks, because the same retrieval process is also needed to refresh the memory traces (Barrouillet *et al.* 2004).

On the other hand, experimental results have shown that attention-demanding processes interfere with each other, suggesting that there is a central attentional bottleneck that allows just one processing step at a time. For example, Garavan (1998) concluded from a study on attention-switching processes that individuals can attend to just one 'object' in working memory at any one time, results that have been replicated and extended by Oberauer (2003). In the same way, Rohrer and Pashler (2003) demonstrated that a secondary task that does not involve verbal material and does not itself rely on retrievals but requires attention (a serial choice-reaction time task) reduces speed and accuracy of recall in a concurrent free recall task. The authors concluded that memory retrieval is subject to severe interference from unrelated central processing. These results are in line with the hypothesis that any processing component that captures attention would impede the refreshment of the decaying memory traces, recall performance depending on the time during which attention is captured. We will see below that our most recent results suggest that the 'capture of attention' hypothesis is probably more plausible than the more restrictive 'bottleneck for retrieval' hypothesis.

Our fourth assumption is that the sharing of attention is achieved through a process of rapid switching of the focus from processing to maintenance and vice versa. However, there is no need to suppose, as Towse *et al.* (1998) did, that this switching only occurs when the participants are presented with a new item to memorize or with a new problem to solve. As demonstrated by Cowan in the field of recall activities (Cowan 1992; Cowan *et al.* 1994), individuals can engage in a covert retrieval process during short pauses to reactivate decaying memory traces. In the same way, it may well be supposed that, while engaged in the counting, operation solving, or reading component of working memory span tasks, participants may keep short pauses free for the brief reactivation of the memory items by means of a simple mental search without engaging in time-consuming activities such as rehearsal (Cowan *et al.* 1994). Thus, we assume that individ-uals cope with the requirements of the dual task by a fast and incessant process of attention switching. There is really a resource-sharing process in working memory functioning, but it is time-based.

Time and cognitive load

This conception of working memory delineates in turn a conception of cognitive load that departs from more traditional conceptions that conflate cognitive load with complexity. As we noted earlier, the processing components of traditional working memory span tasks involve com-plex activities such as reading sentences, counting dots among distractors, solving complex equa-tions or reasoning. The rationale of these tasks is to be found within the theoretical framework of resource-sharing and the hypothesized trade-off between processing and storage (Case 1985): because complex activities are especially resource-demanding, they conflict with the concurrent maintenance of the memory items. There is thus a clear difference between simple span measures assessing short-term memory capacity (e.g., the digit span or the word span), and complex span measures involving storage plus processing that assess working-memory capacity (e.g., counting

or reading spans). Within this framework, the more complex and demanding the processing component, the better and the more reliable the working memory task.

As far as we know, Towse and Hitch (1995) were among the first to cast doubt on this conception by suggesting that the detrimental effect of processing on maintenance was not due to the cognitive load it would involve but to its duration. However, as we have seen, these authors proposed a simple effect of time by assuming that there would be no active maintenance of stored items while the processing component of the task is performed (Towse, Hitch and Hutton 2002). In this case, memory traces would suffer from an uninterrupted decay during the phases devoted to processing, recall performance depending on their total duration. Short processing phases would thus reduce the delay between storage and recall, resulting in better recall performance and higher spans.

The time-based resource-sharing model departs from both previous conceptions. It assumes that the cognitive load that an activity involves does not depend on its complexity per se, but neither does it depend on the total duration of the activity. Those tasks that frequently capture attention for lengthy periods of time would involve a high cognitive load by preventing the possibility of running another activity or maintaining items active in short-term memory concurrently. By contrast, the activities that require short and infrequent periods of attentional focusing would permit the switching of attention to other activities, involving a low cognitive load. According to the time-based resource-sharing model, cognitive load depends on the proportion of time during which a given activity captures attention in such a way that the refreshment of memory traces or any other activity that requires attention is impeded. As far as processing components involved in working memory span tasks are concerned, the higher this proportion, the fewer and shorter the periods during which pauses can be kept free to divert attention from processing to the retrieval and the refreshing of decaying memory traces of the items to be recalled.

It is worth noting that this conception also accounts for the fact that complex activities proved to be highly disruptive on recall when used within working memory span task. As Kahneman (1973) suggested, complex activities are not demanding because they are complex but because time pressure is inherent to their structure. For example, solving the equations involved in Turner and Engle's (1989) operation span task or reading the sentences in Daneman and Carpenter's (1980) reading span task can not be interrupted for long periods of time without loosing the thread of the activity. However, our model predicts that complexity is not needed. As Barrouillet and Camos (2001) suggested, even a simple task that would require a continuous attentional focusing, for example an uninterrupted series of retrievals from long-term memory, should result in a heavy cognitive load and should have a highly detrimental effect on recall when inserted into a working memory span task.

Many of the predictions issuing from our theory have been tested using such tasks in which participants were asked to maintain letters in memory while performing a secondary task consisting of reading series of digits presented one by one on screen. We named this task the reading digit span task. For this kind of task cognitive load – that is, the proportion of time during which attention is captured – would be a function of the number of retrievals to be performed, their nature, and the total time allowed to perform them. Cognitive load (CL) would thus correspond to the following:

$$CL = \Sigma \, aN/T$$

where N corresponds to the number of retrievals (i.e., the number of digits to be read), a to a parameter that represents the difficulty of these retrievals, that is to say the time during which these retrievals totally capture attention, and T to the total duration of the activity (the duration of the interletter interval). In this case, cognitive load can be assimilated to the number of
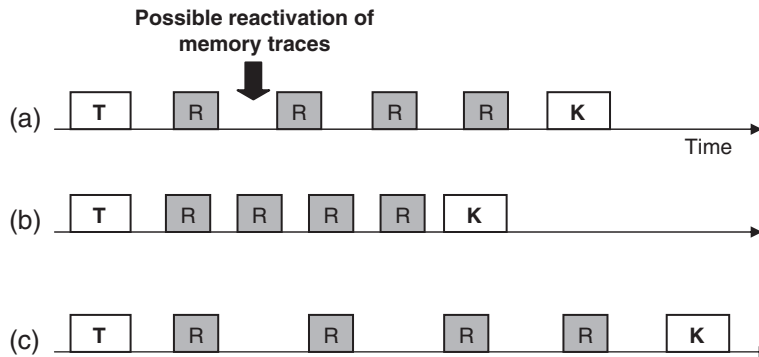
**Possible reactivation of
memory traces**

Figure 4.1 Schematic illustration of the time course of the processes involved in a computer-paced working memory span task in which letters must be remembered while the processing component requires successive retrievals (R). The three panels illustrate how variations in the time allowed to perform a fixed number of retrievals constrain the switching process between processing and maintenance.

retrievals/time ratio. Reading digits is a very simple activity that can be considered as complex. Provided that the digits are presented at a comfortable pace, this activity could allow participants to free up interdigit pauses during which the decaying memory traces of the letters could be retrieved and updated – see panel (a) of Figure 4.1. However, reading the same digits at a faster pace would increase cognitive load by preventing attentional shifting, thus impeding the refreshment of the memory traces and leading to poor recalls. This phenomenon is illustrated in panel (b) of Figure 4.1. By contrast, increasing the time available to read these digits would lead to decrease the cognitive load and would result in better recall performance (panel c).

Two main predictions issue from these proposals. First, working memory spans mainly depend on the rate at which the processing component of the task has to be performed, that is, in our example, on the Number of retrievals/Time ratio. Second, simple activities can have an effect on maintenance which is as detrimental as complex activities provided that they require attention and are also performed at a high rate.

Exploring the cognitive load as the Number of retrievals/Time ratio

The predictions concerning the effect of the Number of Retrievals/Time ratio were tested in a series of experiments using the reading digit span task (Barrouillet *et al.* 2004). The critical aspect of this task lies on the time constraints imposed on the participants, who are not allowed to read the digits at their own pace. In a first experiment (Barrouillet *et al.* 2004, Experiment 4), we manipulated the number of digits to be read while keeping constant the total time allowed to read them. After the presentation of each letter, either 10 or 6 digits were successively displayed on screen at a regular rhythm over a total period of 6 s. In each condition, the length of the series of letters to be remembered was progressively increased from 1 to 7 until the participant failed to recall the letters in the correct order. The time-based resource-sharing model predicts that a larger number of retrievals over a fixed period of time should result in a higher cognitive load and consequently on poorer recall. In line with this hypothesis, the slow-paced condition elicited a higher mean working memory span than the fast-paced condition (4.28 and 2.77 respectively).

In the following experiment, the time allowed to read the digits was manipulated while their number was kept constant (Barrouillet *et al.* 2004, Experiment 5). In the slow-paced condition, participants were given 1s to read each digit, whereas in the fast-paced condition this time was reduced to 600 ms. We predicted that the latter condition would result in a higher cognitive load and thus in poorer recalls. This was exactly what we observed, the slow-paced condition resulting in higher spans than the fast-paced condition (4.76 and 3.01 respectively). This result was of particular interest because it ruled out one of the main predictions issuing from Towse and Hitch's (1995) proposals. Indeed, the fast-paced condition corresponded to a shorter duration of processing. Their theory would predict better recalls because shorter durations of processing involve shorter delays of retention between storage and recall. This was not what we observed, suggesting that, as we surmised, attention must be shared by switching rapidly from processing to maintenance during the intervening activity and that the relative ease of this switching process determines what is called cognitive load.

Finally, we tested the hypothesis that cognitive load is a linear function of the Number of retrievals/Time ratio. Progressively increasing this ratio should result in a smooth and linear decrease in span. For this purpose, nine groups of adult participants were presented with nine different values of the ratio (from 0.4 to 2) resulting from the combination of three different numbers of digits presented (4, 8, or 12) with three total periods (6, 8 or 10 s). As shown in Figure 4.2, recall performance was highly correlated with the ratio, revealing a quasi perfect trade-off between processing
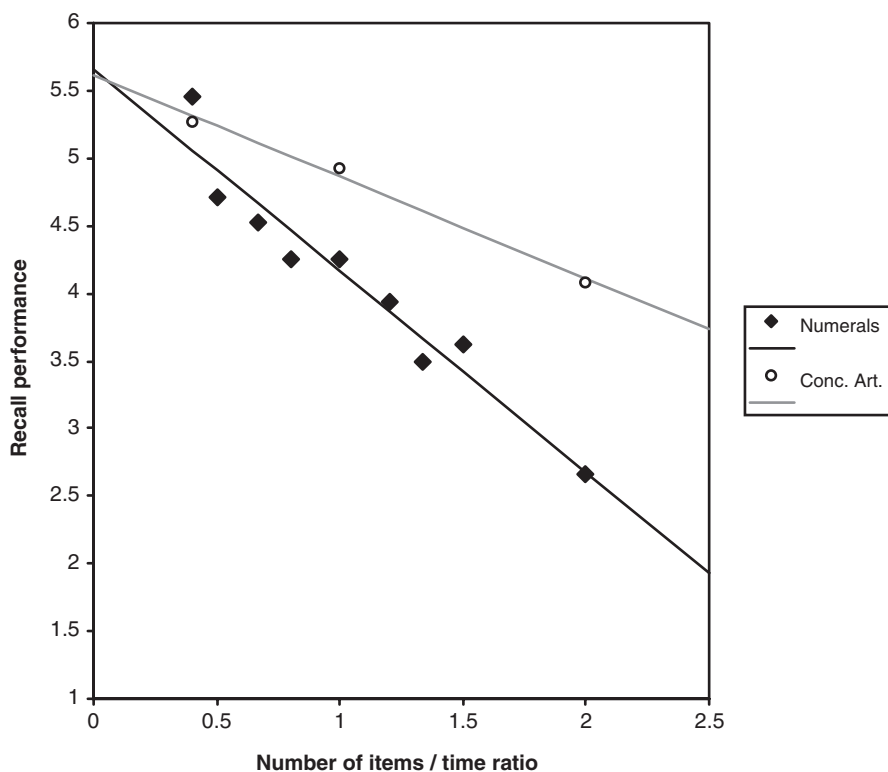

Figure 4.2 Mean working memory spans as a function of the Number of retrievals/Time ratio and the nature of the processing component (reading numerals or saying 'baba') from Barrouillet *et al.* (2004, Experiment 7). Conc. Art. = concurrent articulation.

and storage. It could be argued that this relation was merely due to the different levels of articulatory suppression that the nine ratio values involved and not to any cognitive load. Uttering digits would simply block the articulatory loop and impair the rehearsal of the letters. In order to control for this effect, three additional groups of adults were presented with the same task but the digits were replaced by the syllable BA which participants were asked to read aloud each time it appeared on screen. This control involved only three ratio values of 0.4, 1 and 2 (4, 8 and 12 syllables presented respectively over 10, 8 and 6 s). Our theory predicted that spans should decrease as the ratio increases, even when the task consists only in attending to a signal and always pronouncing the same syllable. However, this decrease should be less pronounced in the control condition than when reading numerals because the participants simply have to keep track of a habituated stimulus and always produce the same response. Accordingly, the resulting spans decreased as the number of syllables to be uttered increased but this effect was smaller than that observed with digits (Figure 4.2).

Lépine, Bernardin and Barrouillet (2005) conducted an even more stringent test of the alternative hypothesis of an effect of the Number of retrievals/Time ratio that would only be due to articulatory suppression. In their Experiment 3, they manipulated the time allowed to perform a task that did not involve any articulatory suppression. As in the previous experiments, participants were presented with a task in which letters were to be remembered while digits were successively presented on screen. However, they were not asked to read these digits but to judge their parity by silently pressing identified keys on keyboard for odd and even. The time allowed to judge the parity of each number was 800 ms in the fast-paced condition and 1500 ms in the slow-paced condition. Once more, the former condition resulted in lower recall performance than the latter (mean spans of 3.81 and 5.10).

Overall, these observations lent strong support to the main assumptions of the time-based resource-sharing model. First, time appears to play a major role in determining the impact of the processing component on the concurrent maintenance of information, but in a subtle manner. For example, reducing the total duration of the processing component did not lead to better but to poorer recalls when the amount of stimuli to be processed is kept constant. Keeping this total duration constant led to poorer recalls when the amount of processing was increased. The linear trend reported in Figure 4.2 made clear that as we hypothesized, what matters is the proportion of time during which the processing component captures attention and impedes the retrieval and refreshment of the decaying memory traces. Second, the smooth decrease in span when the difficulty of the processing component increases suggests that there is a quasi perfect trade-off between processing and storage, exactly as the resource-sharing theoretical framework predicts. Taken together, these two points suggest that the resource-sharing that occurs between processing and storage is time-based.

Though these first results were in line with our theory, several counterintuitive claims still needed to be tested. For example, one of the predictions of our theory is that any simple task that captures attention can disrupt concurrent maintenance as efficiently as a complex activity provided that it has to be performed under sufficient time pressure. The studies reported above demonstrated that time pressure has an effect on recall performance, but they did not directly compare the effect of simple processing under time-pressure with the effect of complex activities such as reading or equation solving. Are simple activities as disruptive as complex tasks used in traditional span tasks? The next section addresses this question.

Is cognitive load a matter of complexity?

Intuitively, reading sentences for comprehension or solving complex equations is more demanding than identifying letters or browsing in the numerical chain by adding or subtracting one.

Indeed, we often experience the former activities as laborious, requiring all our attention, and we are well aware of many possible mistakes. By contrast, we rarely experience the latter activities as arduous but instead as undemanding, nearly automatic, the correct answer coming into our mind without any apparent effort. Cognitive psychology echoes these intuitions based on introspective experiences through the widespread notions of controlled and automatic processes, this opposition underpinning most of the theoretical constructs related to working memory. On the one hand, reading sentences and solving equations undoubtedly pertain to those activities that require selection of relevant information among the flow of incoming stimuli and inner knowledge, to select the appropriate and inhibit the irrelevant strategies, and to control their process, that is activities that are known to involve the central executive and to consume cognitive resources (Engle *et al.* 1999a). On the other hand, identifying letters or browsing in the numerical chain probably rely on direct and automatic retrievals from long-term memory that are often considered as leaving the pool of resources intact (Rosen and Engle 1997). However, according to the time-based resource-sharing model, this distinction would only hold when considered in a temporal vacuum. When taking time into account, reading letters or adding 1 to small numbers could become as demanding and consume the same amount of cognitive resource as reading sentences and solving equations! Should we really believe that the former activities can produce the same disruptive effect on concurrent maintenance in short-term memory as tasks that we experience as being so demanding?

This was tested by Lépine, Bernardin and Barrouillet (2005). The authors compared recall performance in either self-paced traditional working memory span tasks involving complex processing components or new computer-paced tasks involving processing components as simple as those we evoked above. They predicted that simple processing components presented at a slow and comfortable pace should have a low impact on concurrent maintenance, thus resulting in higher spans than the complex processing component involved in traditional working memory span tasks. However, according to the time-based resource-sharing model, they predicted that the difference in spans should disappear when simple processing components have to be performed at a fast pace.

In a first experiment, they used a traditional operation span task inspired from Turner and Engle (1989) in which participants were asked to remember series of letters, each letter being followed by an equation to be verified (e.g., '6 + 7 + 2 = 13?'). The equations remained on screen until the participant pressed one of the two keys identified as 'true' or 'false' on keyboard without any limit of time. This task was compared with the continuous operation span task, first introduced by Barrouillet, Camos and Bernardin (2001; see also Barrouillet *et al.* 2004), in which each equation of the former task was replaced by a root (a number from 1 to 9) followed by a series of sign operand pairs (i.e., + 1 or − 1). Such a series can for example take the following form: 5 / + 1 / + 1 / − 1. Participants were asked to read aloud the root, the sign-operand pairs, and to give all the answers aloud (i.e., 'five, plus one six, plus one seven, minus one six'). In the slow-paced condition, participants had 2 s to process each sign-operand pair, but only 1 s in the fast-paced condition. Many studies in cognitive arithmetic have established that adding or subtracting one to small numbers just involves direct retrievals of the answer from memory, retrievals that have been described as automatic (Aschcraft and Battaglia 1978; Barrouillet and Fayol 1998; LeFevre, Bisanz and Mrkonjic 1988). The authors were thus not surprised to observe that the slow-paced condition resulted in higher spans than the traditional operation span task (3.65 and 2.30 respectively). However, and as the time-based resource-sharing model predicted, when the participants were subjected to a fast pace, their continuous operation span dramatically dropped to the same level as the operation span (Table 4.1).

However, the authors acknowledged that this result was not so surprising. As we stated above, the cognitive demand of complex activities such as equation solving lies on the time pressure

Table 4.1 Mean spans (and standard deviations) as a function of the nature of the processing component involved in the working memory span tasks

	Nature of the processing component		
	Simple computer-paced tasks		**Complex self-paced tasks**
	Slow	Fast	
Operation solving	3.65	2.30	2.30
	(0.92)	(0.78)	(0.81)
Reading	4.17	3.11	3.38
	(1.00)	(0.93)	(0.78)

Adapted with permission from Lépine, R, Bernardin, S and Barrouillet, P (2005). Attention switching and working memory spans. *European Journal of Cognitive Psychology, 17*, 329–346.

inherent to their structure. The authors noted that mental arithmetic necessitates keeping track of the problem to be solved, of subgoals and of intermediary results, the memory traces of which suffer from decay and interference. Thus, any interruption or slowing down calculation can lead to irremediable loss of information or failure. Because it could be argued that the processes involved in equation solving are akin to those involved in solving the continuous operations, what this experiment demonstrated was that a time pressure induced by the experimental design has the same effect as the time pressure inherent to the structure of complex activities. Moreover, the processing component of the continuous operation span task involves a memory load because the participant must keep track of the current state of the calculation, a memory load that could have impaired the concurrent maintenance of the letters. Thus, the authors decided to compare the effects on span of two activities that deeply differ in the cognitive process they involve.

In a further experiment, they compared a reading span task inspired from Daneman and Carpenter (1980) in which participants were asked to read aloud and evaluate the plausibility of sentences displayed on screen (e.g., 'A cow lays eggs') with a reading letter span task in which the processing component consisted in reading series of letters displayed successively on screen at a fixed pace. In both tasks, each sentence or series of letters was preceded by a to-be-remembered number presented on screen. Of course, identifying letters is one of the elementary components of reading, but reading series of letters even under time pressure can not be considered as mimicking the activity of reading comprehension. Analyzing the latter activity, Siegel (1994) pointed out that triggering grapheme–phoneme conversion rules, retrieving information about word meaning and processing syntax are cognitive processes involving executive components, all activities that are obviously not needed to read letters. Nonetheless, the authors made the same predictions as in the first experiment because reading letters require attention and a sufficiently fast-paced presentation would prevent attentional switching. In line with this prediction, they observed that the traditional reading span was lower than the reading letter span only when letters were presented at a slow rate (one letter every 1300 ms), but this difference disappeared when letters were presented at a rate of one letter every 675 ms (Table 4.1).

This last result demonstrated that an activity that does not require any memory load or algorithmic process but only the retrieval of overlearned information from memory such as reading letters is sufficient to disrupt concurrent maintenance. As Barrouillet and Camos (2001) surmised, even a fairly simple task that continuously captures attention has the same detrimental effect on span as the complex activities involved in the traditional working memory span tasks. Thus, even if complex activities often involve a high cognitive load, cognitive load is not a matter of complexity. Lépine, Bernardin and Barrouillet (2005) noted that it could be considered as

surprising that the simple activities they used in their computer-paced tasks have so great an effect on concurrent maintenance because these activities have been considered as automatic and non-demanding. It has often been suggested that solving simple additions and subtractions used in the continuous operation span task would rely on a process of direct and automatic retrieval of the answer from memory (LeFevre *et al.* 1988; Winkelman and Schmidt 1974; Zbrodoff and Logan 1986). However, they also noted that it had already been pointed out by Kahneman (1973) that a conception of the cognitive demand conflated with complexity was probably wrong. Kahneman reported that when measuring mental effort by arousal and pupillary dilations, easy tasks like the recall of thoroughly overlearned information or retaining five digits for immediate recall induced larger pupillary dilation than apparently more complex activity. More importantly, Kahneman noted that the amount of effort that is required to perform a task could not merely depend on intrinsic characteristics of this task because it is also obviously determined by the rate at which it is performed. The present results confirmed this supposition: even simple tasks can become as, and probably more, demanding than activities considered as complex.

Beyond predicting this counterintuitive result, the time-based resource-sharing model accounts for the fact that we do not usually experience simple activities as demanding. Remember that what matters is time. One of the main properties of the simple tasks we use (identify a letter, finding the answer of 4 + 1) is to extend over very short periods of time. Thus, they can easily be inserted in any activity without interrupting the thread of one's thoughts. These activities capture our attention for such short periods of time that, most of the time, we fail to notice them. They are in fact so fast that we have the impression of performing them in parallel with other more complex activities that obviously solicit our attentional capacities. However, this is an illusion, as their effect on concurrent activities when performed under time pressure made clear.

Although this study demonstrated that our computer-paced tasks can become as difficult as traditional working memory span tasks involving complex activities, an important issue concerning these tasks remained to be addressed. The main interest of the complex span tasks is not that they are more difficult than simple span tasks, or that they allow us to improve our understanding of working memory functioning, but that they provide us with spans that are good predictors of high-level cognition and academic achievement. However, do the spans provided by our new tasks have the same properties as the traditional reading or operation spans? After all, it remains possible that, contrary to the complex activities of reading sentences and solving equations, the simple activities we use disrupt maintenance by affecting peripheral and unimportant processes. In this case, the spans collected through the new tasks would not reflect individual's cognitive capacities and would not provide valid measure of working memory capacity, validity being judged by the correlation between the measure and higher level measures of cognition (Engle, Tuholski, Laughlin *et al.* 1999). In other words, our tasks could lack the predictive power that makes working memory span tasks important for psychology.

Are the new time-constrained tasks as predictive of high level cognition as the traditional working memory span tasks?

This question was addressed by Lépine, Barrouillet and Camos (2005). The authors noted that two alternative hypotheses could be put forward to account for the well-known relationship between working memory spans and performance on high level cognitive activities. According to the first hypothesis, traditional working memory tasks would evaluate some general cognitive capacity that our model describes as an attentional capacity involved in any cognitive process requiring access and maintenance of items of knowledge. This limited attentional capacity would thus underpin and constrain most of the processing steps of the complex activities involved in

high-level cognition, such as reasoning or problem-solving. According to this view, complexity is an unnecessary characteristic of the processing component of a valid working memory task because what is required is an activity that solicits and captures attention. Moreover, the temporal constraints of our computer-paced tasks hamper the use of possible strategies for coping with the demands of the dual-task paradigm. It is worth to note that these strategies are allowed by most of the traditional self-paced working memory tasks in which participants are free to interrupt and resume their activity as they wish (Baddeley, Logie, Nimmo-Smith *et al.* 1985; Case, Kurland and Goldberg 1982; Daneman and Carpenter 1980; Turner and Engle 1989). According to the attentional capacity hypothesis, these strategies may produce biased measures of the fundamental capacity by allowing the most skillful participants to overcome their cognitive limits. Thus, this hypothesis predicts that the spans provided by the new tasks will be more predictive of performance in high-level cognition than the traditional tasks.

An alternative hypothesis would be that the ability to plan and use the strategies described above is the basis of the relation between working memory spans and high-level performance. High working memory span individuals would be those who are better able to plan and monitor their activity in complex situations strategically, achieving better performance in working memory dual tasks and in higher-level cognition. In this account, working memory spans are predictive because the traditional tasks mimic high-level cognitive activities, which require the individual to simultaneously maintain goals and intermediary results and to run complex operations. This hypothesis predicts that the traditional spans will have a greater predictive value because the new tasks involve less strategic factors and involve only elementary processes.

The authors tested these two hypotheses by comparing the correlations between academic achievement of 11-year-old children on the one hand, and working memory spans evaluated either by traditional or new span tasks on the other. The tasks were the same as those used by Lépine, Bernardin and Barrouillet (2005) and described above: that is a reading and an operation span task for the traditional tasks and the reading letter and the continuous operation span task for the new tasks. The level of academic achievement was provided by individual scores from the national academic achievement test that each French sixth grader takes at the beginning of the academic year. This test gives compound scores in literacy and mathematics, as well as a global scholastic score in terms of percentage of success. The results were particularly clear and eloquent (Table 4.2).

Table 4.2 Correlations between the traditional and the new working memory span tasks on the one hand and the scholastic scores on the other for 93 French sixth graders

Working memory tasks	Scholastic score		
	Literacy	**Mathematics**	**Global**
Traditional			
Reading span	0.30	0.33	0.34
Operation Span	0.32	0.35	0.36
Traditional compound score	*0.34*	*0.38*	*0.39*
Computer-paced			
Reading letter span	0.46	0.48	0.50
Continuous operation span	0.39	0.41	0.42
New compound score	0.50	0.52	0.54

Adapted from Lépine, R, Barrouillet, P and Camos, V (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin and Review, 12*, 165–170, Copyright Psychonomic Society Inc.

Though the traditional reading and operation span were good predictors of academic achieve-ment, both new tasks were better correlated with each of the subcomponents as well as with the global score of the scholastic evaluation. Stepwise regression analyses revealed that the new working memory spans were better predictors of the scholastic scores in mathematics and literacy than the traditional spans, the residual part of variance accounted for by the traditional spans never reaching significance.

The authors concluded that the complexity of the processing component in most of the tradi-tional working memory span tasks is a superfluous characteristic. As we noted above, self-paced working memory span tasks require complex activities to induce the necessary time pressure that is inherent to their structure. However, as soon as the tasks are not self-paced, attentional shifting is strongly constrained and complexity becomes unnecessary. As a consequence, the predictive value of the traditional spans does not stem from their capacity to assess an ability to cope strate-gically with the demands of complex span tasks. Indeed, when the possibility of dealing strategi-cally with the task is reduced by computer-paced presentation, the predictive value is increased. Rather, working memory span tasks measure a fundamental capacity related to the attention required for retrievals from memory during reading of letters or solution of the simple continu-ous operations, but also for reactivation of the decaying memory traces. Thus, those tasks that strongly constrain free attention shifting from processing to maintenance and permit better control of the strategies that undermine the rationale of working memory span tasks provide better estimates of individuals' amount of attention, that is of their working memory capacities.

Discussing the interest of their new working memory tasks, Lépine, Barrouillet and Camos (2005) suggested that another source of the higher predictive power of these tasks could stem from the probable lower interindividual variability in simple processes. Individuals can differ greatly in literacy and mathematics, and these differences in processing efficiency could under-mine a reliable measure of their real cognitive capacities. By contrast, reading letters and brows-ing the numerical chain are universal skills in the literate population, probably with low differences in proficiency. However, does this mean that any simple activity could be used as a processing component within a working memory span task and provide us with a valid measure of working memory capacity? Does any activity that captures attention elicit a cognitive load dis-rupting maintenance in such a way that the recall performance reflects working memory capaci-ties? Answering these questions requires clarifying the mechanisms by which the processing component impairs concurrent maintenance in short-term memory.

Does the nature of the activity matter, or just the capture of attention?

The previous text established that complexity is not needed to produce cognitive load. Does this mean that the nature of the activity does not matter at all? Is cognitive load simply created by the attentional capture? And what characteristic of this capture is important: its intensity or its dura-tion, or both? Unfortunately, we can not yet answer these questions. What we can do is to suggest some leads.

In order to clearly establish the role of attention in the effect of the processing component, Gavens and Barrouillet (2004) compared two tasks that differed only by a small but crucial detail. They had children to perform a reading digit span task as we described above. In one condition, as in the usual design of this task, the digits to be read between the letters appeared at random, but on the other condition, the same digits were ordered: instead of being presented with series of the form $6 - 3 - 5 - 2 - 4 - 1$, the children were presented with the canonical order $1 - 2 - 3 - 4 - 5 - 6$. Apart from this difference in order of presentation, both conditions presented exactly

the same items, required the same activity of reading, and involved the same total duration of processing and the same level of articulatory suppression because the same words were uttered. Nonetheless, the authors predicted that the ordered condition should induce higher recall performance. They reasoned that an ordered presentation would reduce the attentional demand of the task by making it possible to anticipate the nature of the following stimulus and preactivating its phonological representation. By contrast, a random presentation would not only disallow this possibility but would even require to inhibit interferences that could result from spreading activation within the numerical chain. As the authors predicted, the random presentation resulted in lower spans than the ordered presentation in two different groups of children aged 8 and 10 respectively. This experiment clearly demonstrated that the cognitive load of an activity does not depend only on its nature or even on the rate at which it is performed, but that subtle changes on the attentional demand it involves could have substantial effects on a concurrent task. It is interesting to note that all the experiments we ran in our laboratory in order to extend these results to adult participants failed: the difference between the random and ordered conditions disappears with age. This is not perhaps surprising as we can surmise that the access to the phonological representation of digits is highly automatized in adults who no longer take advantage from the ordered presentation.

Although Gavens and Barrouillet's (2004) results indicate that cognitive load depends in some way on attention, they do not provide any evidence that what matters is the duration of this capture. Unfortunately, we did not register the response times during the completion of the task. According to the time-based resource-sharing model, it should be the case that reading times in children are longer for the random than the ordered presentation and that this difference disappears in adults, but we have not yet verified this point. However, Barrouillet, Bernardin, Portrat *et al.* (2007) provided evidence that what matters in the reading digit span is the reading time, and thus probably the duration of the capture of attention. In two experiments, they compared the reading digit span from tasks in which the format of the digits was manipulated. Digits were presented on screen either in their verbal (e.g., 'four', 'six'), Arabic (i.e., 4, 6), or Roman (IV, VI) form. The authors verified in a pretest that the Roman format elicits longer reading times (a mean of 625 ms in adults) than both the verbal and the Arabic forms (446 ms and 442 ms respectively). It can thus be assumed that reading Roman digits captures attention over longer periods of time than reading the same digits written in their Arabic or verbal forms. Thus, they predicted that, when inserted in a reading digit span task, the Roman digits should have a more detrimental effect on span than the verbal and Arabic digits. This is exactly what they observed. The Roman condition in which the digits took longer to be read resulted in lower mean spans (3.87) than the verbal and Arabic conditions, which did not differ in reading times and accordingly resulted in the same level of recall performance (mean spans of 4.50 and 4.54 respectively). In a subsequent experiment, Barrouillet *et al.* (2007) extended this result to a non symbolic representation of digits by using dice-like canonical patterns of 1 to 6 dots. A pretest revealed that identifying quantities of dots (507 ms) took longer than reading the equivalent numbers in their Arabic or verbal form (424 ms and 425 ms respectively). When interleaved into series of letters to be remembered, these different stimuli resulted in recall performance that depended on reading times (75 per cent for dots compared to 82 per cent for both Arabic and verbal forms). Barrouillet *et al.* (2007) observations strongly suggest that in working memory tasks such as the reading digit span task, the detrimental effect on span of the processing component is closely related to the time during which it captures attention. Even small differences can result in significant changes in the amount of information that can be concurrently maintained in short-term memory and recalled. But does the nature of the activity that captures attention matter? Remember that, when presenting the time-based resource-sharing model, Barrouillet *et al.* (2004) suggested that some

activities could be more damaging than others for the concurrent maintenance of information in short-term memory.

The effect of the Number of retrievals/Time ratio as well as the results reported in this section suggest that time factors play a major role in the conflict that occurs between the processing and maintenance components of the memory span tasks. According to our theory, this is due to the fact that maintaining items in short-term memory requires their frequent refreshment by a process of retrieval that necessitates attentional focusing (Cowan 1999). However, the locus of the conflict could lie either on the occupation of the process of retrieval or on the capture of attention. According to the first hypothesis, the processing component impedes maintenance as far as it occupies the retrieval process, which is thus unavailable for refreshing the decaying memory traces. As a consequence, the critical factor should be the number and duration of the retrievals that the processing component involves, and secondary tasks necessitating frequent retrievals from memory should have the most detrimental effect on working memory span tasks. According to the second hypothesis, the processing component disrupts maintenance because it blocks the focus of attention. Thus, the critical factor should be here the duration of this capture, whatever the nature of the activity that produces it. As a consequence, activities differing in nature but occupying the focus of attention over equivalent periods of time should have the same effect on concurrent maintenance.

A first investigation of this problem was conducted by Bernardin, Portrat and Barrouillet (2006) who compared the effect on maintenance of two secondary tasks that involved exactly the same stimuli but differed on the processes they involved, depending on whether they needed to retrieve information from memory or not. In both conditions, participants were presented, as in the previous studies, with series of letters to be remembered. After each letter, 7 numbers from 1 to 10 were successively presented on screen over a total period of 9 s. Each of these numbers was displayed either above or below a horizontal line centered on screen. In one condition, which was assumed to involve retrievals, the participants were asked to judge the parity of the numbers by responding either 'odd' or 'even'. In the other condition, they had just to judge their spatial location by responding either 'high' or 'low' when the number was above or below the horizontal line respectively. This latter activity involves a response selection, as in the parity judgment task, but no retrieval from long-term memory. Apart from this difference, both conditions involved exactly the same stimuli presented at the same rate. The results were particularly clear, the spatial task eliciting higher spans than the parity task (mean spans of 4.43 and 3.35 respectively). At a first glance, this difference supports the 'bottleneck for retrieval' hypothesis. A secondary task that calls for retrievals from memory has a more disruptive effect on concurrent maintenance than a task that certainly requires central processes (e.g., response selection), but not retrievals. However, this difference in span is far from being compelling. Indeed, as a posttest using a vocal key made clear, adults are faster in giving their answers to the spatial than the parity task (411 ms vs 560 ms). Thus, the difference in span could also result from differences in the time during which the two tasks captured attention in our participants. The longer this capture, the lower the spans. In order to disentangle the two competing hypothesis, an experiment that would contrast different secondary tasks the duration of which is carefully controlled was needed. We recently ran such an experiment in our lab.

Barrouillet *et al.* (2007) used the same working memory span tasks involving either spatial or parity judgment as in Bernardin *et al.* (2006), except that the participants gave their responses by pressing keys on keyboard. This modification made it possible to easily measure the time that the participants allocated to the secondary task by registering their response times for each stimulus to be processed. Moreover, following the same logic as in many of our previous studies, we varied the rate at which the numbers were presented between the letters. Each letter was followed by

series of either 4, 6 or 8 numbers over a fixed total period of 6400 ms. Suppose that the nature of the activity does not matter at all, the time during which attention is captured being the sole and unique factor in accounting for memory loss in working memory span tasks. In this case, as Bernardin *et al.* (2006) observed, the spatial judgment task would result in higher recall performance than the parity judgment task, but when the two tasks would involve the same duration of attentional capture, these differences should disappear. Of course, we can not measure this duration precisely because it is evident, as pointed out by Rohrer and Pashler (2003), that the centrally demanding stages of an activity do not occupy no more than a sizable fraction of the time consumed by this activity. Nonetheless, we assumed that two activities that consume the same time could be considered as capturing attention in approximately the same way. Thus, in each experimental condition, we computed the actual processing time within the interletter intervals by adding the reaction times to the 4, 6, or 8 stimuli that were processed after each letter (this actual processing time is noted $\sum$ RT). Obviously, this $\sum$ RT increased with the number of stimuli to be processed and, as we previously observed, the parity judgments elicited longer processing times than the spatial judgments. In line with our previous results, the mean spans decreased as the rate at which the numbers were processed increased, and the spatial task resulted in higher spans than the parity task (Table 4.3). However, our main interest was not in these expected and already known phenomena but on the specific effect of time and tasks on spans. Of course, as we noted, an analysis of variance with the type of task and the number of items presented as independent variables revealed highly significant effects on spans for the two factors without any interaction. However, when the individual mean $\sum$ RT were introduced as covariate in an analysis of covariance, time appeared as a statistically significant covariate whereas the effect of tasks on spans became non significant, the F value dramatically dropping from 19.64 to 0.10. This phenomenon can be illustrated by plotting the observed spans with the $\sum$ RT for each task and by linearly extrapolating what would have been, for example, the recall performance from the spatial judgment task if it had consumed the same time as the parity judgment task (Figure 4.3). As we have seen, the two tasks resulted in quite different overall mean spans (4.48 for the parity and 5.23 for the spatial judgment tasks), but this difference disappeared when imagining that the two tasks would consume the same time.

The main conclusion that can be drawn from these experiments is that the relevant factor in determining the cognitive load induced by a given activity is time, whatever the nature of this activity. Though memory retrievals are needed to refresh the decaying memory traces in short-term

Table 4.3 Mean actual processing times and spans as a function of the nature of the processing component and the number of stimuli to be processed.

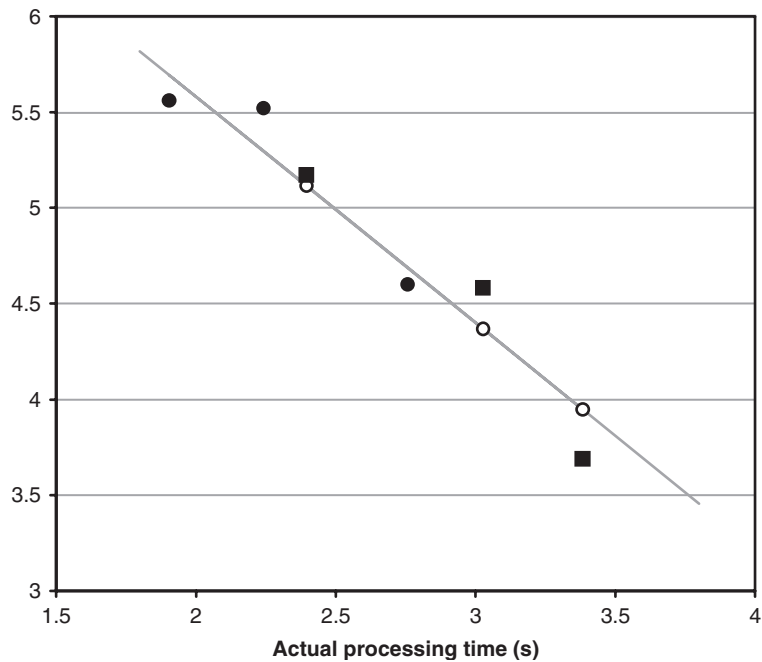| Number of stimuli	Nature of the Processing component			
	Spatial judgment		Parity judgment	
	Actual proc. time (ms)	Span	Actual proc. time (ms)	Span
4	1928	5.56	2467	5.16
	(233)	(0.75)	(400)	(0.78)
6	2297	5.52	3251	4.58
	(239)	(0.62)	(316)	(1.23)
8	2827	4.60	3724	3.69
	(266)	(0.82)	(218)	(0.63)

Figure 4.3 Mean spans as a function of the actual time consumed by the processing episodes for a task of spatial judgment (black circles) and of parity judgment (black squares). White circles refer to linearly extrapolated values if spatial judgments had consumed the same time as parity judgments.

memory, secondary tasks that involve frequent retrievals do not disrupt maintenance in a greater extent than other tasks involving other central processes such as response selection. By contrast, the longer the time consumed to perform this task, the lower the resulting span. Of course, these studies only provide us with preliminary results that deserve replication and extension. However, they are sufficient to rule out the restrictive 'bottleneck for retrieval' hypothesis. This is not to say that there is no bottleneck for retrieval, but instead that this bottleneck seems to be a central bottleneck that affects retrievals, but also other central processes. Our results suggest that, in working memory span tasks, recall performance is a function of the actual time during which this central bottleneck is occupied by the processing component, thus lending strong support to one of the main assumptions of the time-based resource-sharing model.

Conclusions and perspectives

The studies reported in this chapter have many implications concerning the way we consider working memory functioning as well as our conceptions about cognitive resources, resource sharing, and cognitive load. Ever since the seminal studies of Moray (1967) and Kahneman (1973) on attention and mental effort, many models have adopted the capacity or resource theory (Navon 1984) and assumed that there is a general limit on the human capacity to perform mental work. Based on a crude physical analogy, this theory posits that mental activities require the allocation of a certain amount of effort and consume energy (Pascual Leone 1970). Difficult tasks would demand more effort, and thus require a greater amount of mental energy. Because the total amount of energy, the total processing space, or the mental power that can be supplied is limited, performance deteriorates when the demands exceed the available capacities or when a

part of the limited pool of energy is reallocated from the current task to a concurrent and demanding activity (Case 1985). In line with this seminal conception, most of the current models of working memory assume that there is a limited-capacity central executive involved in all the activities that require controlled attention. This kind of attention would be required in order to select the relevant information from the flow of incoming stimuli and inner knowledge, to select the appropriate and inhibit the irrelevant strategies, and to control their processing (Baddeley 1996; Engle *et al*. 1999a; Norman and Shallice 1986). Other models consider attention to be a kind of mental energy that produces activation and determines which information enters working memory and then becomes available for processing (Cowan 1995; Just and Carpenter 1992; Lovett *et al*. 1999). Despite their discrepancies, all these models share the same basic tenets of resource theories: performance depends both on the amount of resources allocated and on the intrinsic demands of the task, with more difficult and demanding tasks resulting in poorer performance. Thus, the simultaneous completion of two demanding tasks results in a trade-off phenomenon because at any point in time, the cognitive resource is shared between the two competing activities: when resources are reallocated from a task A to a task B, performance on task A deteriorates. Thus, common sense, introspective scrutiny as well as elaborated scientific theories have converged in considering that complex cognitive activities such as reading comprehension, mental calculation, or reasoning on abstract material are among the most resource demanding. Should we abandon these ideas? Certainly not. Our minds are the best witnesses of what is hard for them, and apart from rare and famous prodigies, nobody finds easy to calculate the square of 647 in one's head, apparently for good reasons. Accordingly, when involved in dual-task paradigms, these complex activities proved highly damaging for concurrent processes, suggesting that they annex a substantial part of the available resources, whatever the nature of these resources and the way they are shared.

However, what we demonstrated above is that although complex tasks are the most effective in disturbing concurrent maintenance, complexity is not needed and very simple processes, often described as automatic, can have a highly detrimental effect on spans. Thus, it appeared that the relevant factor is not complexity per se but time! What kind of working memory theory could accommodate this unexpected fact? We claim that the time hypothesis constitutes the simplest way to account for trade-off phenomena, cognitive load, and working memory functioning while retaining the main assumptions of cognitive psychology and even of resource theory.

According to Kahneman (1973), work is done and effort is mobilized in response to the changing demands of the task in which the subject engages. This conception echoes current models of working memory that assume that the role of attention is to update the current content of working memory. This updating would involve encoding relevant stimuli, retrieving items of knowledge from long-term memory, and keeping these items active (Anderson and Lebiere 1998; Cowan 1995). Switching attention from one item to another, selecting among activated responses, and maybe inhibiting irrelevant information are processes that could be added to this list. All these activities, usually considered as executive functions, require the allocation of attention controlled by some central executive (Baddeley and Logie 1999; Cowan 1995, 1999, Engle *et al*. 1999a). In other words, we assume that work is done and effort is mobilized each time the content of working memory is modified by controlled attention, that is each time executive functions are involved.

What are the limits of such a working memory? We assume that there are two main limits, both related to time. The first concerns the activation of those items that are outside the focus of attention because this activation suffers from a time-related decay. Note that this phenomenon would not be a limit per se if it was possible to retrieve these decaying memory traces and thus refresh them through a process of redintegration (Hulme *et al*. 1997) while simultaneously performing other attention demanding operations. However, and this is the second limit, it seems that executive

processes are constrained by a central bottleneck and are thus sequential in nature. As a consequence, keeping active the current goals and the information relevant for the activity in hand necessitates frequent interruptions of this activity. When the impairment of maintenance in short-term memory and the further recall performance are taken as a measure of the cognitive load of a given task performed concurrently, this cognitive load is a function of the proportion of time during which this task occupies the central bottleneck and impedes other executive processes to be performed. Thus, as we observed, simple tasks that nonetheless involve executive processes can become highly demanding when performed under time pressure. Moreover, the nature of the executive process that temporarily occupies the central bottleneck does not matter. Whatever this process, it impedes other processes to be concurrently run. As a consequence, what matters is the actual processing time.

Although the time-based resource-sharing model appears to be the opposite of the resource theory exposed above, it is in fact related to this general theory in at least two ways. First, when considering working memory functioning at the atomistic level of the successive elementary processing steps, it is true that our theory contradicts the assumption of a continuous sharing of the resource at any point in time. However, at a macro level of analysis, we can still describe the phenomena we reported in this chapter as resulting from a resource sharing between tasks 'simultaneously' performed. The quasi-perfect trade-off between processing and storage we reported in Figure 4.2 could still be described as a decrease in recall performance while the 'difficulty' of the secondary task increases. Indeed, even if the effect on spans of apparently simple tasks such as reading digits or continuous operations can surprise, those participants who were subjected to fast paces frequently reported the extreme difficulty of these tasks and their firm intention never to take part in working memory experiments! Thus, our estimation of the cognitive load as a time ratio actually reflects conscious experience of mental effort. At this macro level, our results take further and reinforce the physical analogy on which the resource theory is based. Mental activities not only consume energy, but this consumption seems to conform to a general law of physics. We have seen that the recall performance smoothly decreased as the Number of retrievals/Time ratio increased suggesting that the amount of 'resources' consumed by reading digits was a function of this ratio. In other words, the mental power needed to perform a given task corresponds to the rate at which energy is converted, i.e. the amount of work this task demands, here the number of retrievals to be performed, divided by the time taken to perform it. This corresponds to the physical law of power: Power = Work/Time.

Second, at a micro level of analysis, we have assumed that work is done each time executive functions are involved. Let us propose that the energy, the resource that is converted by executive processes is attention. For example, within the ACT-R framework (Anderson 1993; Anderson and Lebière 1998) as well as in Cowan's theory, it is explicitly assumed that attention produces activation and triggers retrievals. Higher amounts of attention result in higher levels of activation. Thus, individual and developmental differences in the amount of attention would account for differences in working memory capacity. Indeed, a higher amount of attentional resources would result in faster executive processes such as retrievals, response selection, updating, and in turn in shorter periods of occupation of the central bottleneck, but also on higher levels of activation and thus on longer periods before the total disappearance of items maintained in short-term memory. Accordingly, Gavens and Barrouillet (2004) demonstrated that even when the difficulty of processing within a continuous operation span task is equated across age, working memory spans still exhibit a developmental increase. The authors suggested that this increase could be underpinned by a greater amount of cognitive resources in older children. Thus, even if the main phenomenon in working memory is probably time sharing, this does not necessarily make the notion of resource useless.

In a fascinating article, Navon (1984) wondered if 'resources' was a theoretical soup stone. The article began with the adaptation of a Russian folktale in which a scamp taught a fool how to prepare a delicious soup that requires nothing but a soup stone. All one has to do is to put the soup stone into boiling water, but adding some vegetables, meat, salt and pepper would not harm the soup and would even achieve a better taste. Navon suggested that if the concept of resource was not sufficient to impose conceptual organization on phenomena in the field, then it might be suspected that that concept is actually a theoretical soup stone that can be taken out, the soup being as good as it was. There are so many issues related to working memory that remain unsolved, and the studies that we reported in this chapter are so limited in scope that we do not know if resources is a theoretical soup stone, but what is clear is that it is difficult and probably premature to take the hot stone out of the soup. Indeed, as we have seen, it remains difficult to account for working memory phenomena without any recourse to the notion of resource. Of course, time plays the major role and central processes seem to use in turn a commodity that can not be shared, but what is this commodity needed by retrievals, response selection and probably other processes? Many models of working memory suggest that it corresponds to controlled attention, the available amount of which could vary from one individual or one age to another, exactly as a resource. Thus, in the present state of our knowledge, the most accurate description of working memory functioning that we can propose corresponds to a time-based resource-sharing.

References

Anderson, JR (1993). *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

Anderson, JR and Lebiere, C (1998). *Atomic Components of Thought*. Hillsdale, NJ: Erlbaum.

Aschcraft, MH and Battaglia, J (1978). Cognitive arithmetics: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 527–538.

Baddeley, AD (1990). *Human Memory*, *Theory and Practice*. Hillsdale, NJ: Erlbaum.

Baddeley, AD (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, *49A*, 5–28.

Baddeley, AD (2000). Short-term and working memory. In E Tulving and FIM Craik (eds), *The Oxford Handbook of Memory* (pp. 77–92). Oxford: University Press.

Baddeley, AD and Hitch, G (1974). Working memory. In GA Bower (ed), *Recent Advances in Learning and Motivation*, Vol. 8 (pp. 647–667). New York: Academic Press.

Baddeley, AD and Logie, RH (1999). Working memory: The multiple-component model. In A Miyake and P Shah (eds), *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge: Cambridge University Press.

Baddeley, AD, Logie, R, Nimmo-Smith, I and Brereton, N (1985). Components of fluent reading. *Journal of Memory and Language*, *24*(1), 119–131.

Barrouillet, P, Bernardin, S and 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 and Camos, V (2007). Time and cognitive load in working memory. *Manuscript under revision*.

Barrouillet, P and Camos, V (2001). Developmental increase in working memory span: Resource sharing or temporal decay ? *Journal of Memory and Language*, *45*, 1–20.

Barrouillet, P, Camos, V and Bernardin, S (2001). *Temporal decay and resource sharing in working memory span measures: A new paradigm*. Paper presented at the third International Conference on Memory, 16–20 July, Valencia (Spain).

Barrouillet, P and Fayol, M (1998). From algorithmic computing to direct retrieval: Evidence from number- and alphabetic-arithmetic in children and adults. *Memory and Cognition*, *26*, 355–368.

Bernardin, S (2004). Mémoire de travail et contraintes cognitives: Le modèle de partage temporel des ressources. Unpublished Doctoral Dissertation. Dijon: Université de Bourgogne.

Bernardin, S, Portrat, S and Barrouillet, P (2006). Impact de la nature et du rythme de realisation de la tâche secondaire dans les tâches d'empan de mémoire de travail. *L'Année Psychologique*, *106*(1), 23–42.

Cantor, J and Engle, RW (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1101–1114.

Case, R (1985). *Intellectual Development: Birth to adulthood*. New York: Academic Press.

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

Conway, ARA and Engle, RW (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *4*, 354–373.

Conway, ARA, Kane, MJ and Engle, RW (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552.

Cowan, N (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, *31*, 668–684.

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

Cowan, N (1999). An embedded-process model of working memory. In A Miyake and P Shah (eds), *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.

Cowan, N, Keller, TA, Hulme, C, Roodenrys, S, McDougall, S and Rack, J (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory and Language*, *33*, 234–250.

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

Engle, RW, Cantor, J and Carullo, JJ (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *5*, 972–992.

Engle, RW and Oransky, N (1999). The evolution from short-term memory to working memory: Multi-store to dynamic models of temporary storage. In RJ Sternberg (ed), *The Nature of Cognition* (pp. 515–556). Cambridge, MA: MIT Press.

Engle, RW, Kane, MJ and Tuholski, SW (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A Miyake and P Shah (eds), *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge: Cambridge University Press.

Engle, RW, Tuholski, SW, Laughlin, JE and Conway, ARA (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331.

Garavan, H (1998). Serial attention within working memory. *Memory and Cognition*, *26*(2), 263–276.

Gavens, N and Barrouillet, P (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory and Language*, *51*, 644–657.

Halford, GS, Maybery, MT, O'Hare, AW and Grant, P (1994). The development of memory and processing capacity. *Child Development*, *65*(5), 1338–1356.

Hitch, G, Towse, JN and Hutton, U (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, *130*(2), 184–198.

Hulme, C, Roodenrys, S, Schweickert, R and Brown, GDA (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1217–1232.

Just, MA and Carpenter, PA (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

Kahneman, D (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

LeFevre, J, Bisanz, J and Mrkonjic, L (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. *Memory and Cognition*, *16*, 45–53.

Lépine, R, Barrouillet, P and Camos, V (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin and Review*, *12*, 165–170.

Lépine, R, Bernardin, S and Barrouillet, P (2005). Attention switching and working memory spans. *European Journal of Cognitive Psychology*, *17*, 329–346.

Lovett, MC, Reder, LM and Lebière, C (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In A Miyake and P Shah (eds), *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 135–182). Cambridge: Cambridge University Press.

Moray, N (1967). Where is capacity limited? A survey and model. *Acta Psychologica*, *27*, 84–92.

Navon, D (1984). Resources – a theoretical soup stone? *Psychological Review*, *91*(2), 216–234.

Norman, DA and Shallice, T (1986). Attention to action: Willed and automatic control of behavior. In RJ Davidson, GE Schwartz and D Shapiro (eds), *Consciousness and Self-regulation*, Vol. 4 (pp. 1–18). New York: Plenum.

Oberauer, K (2003). Selective attention to elements in working memory. *Experimental Psychology*, *50*(4), 257–269.

Pascual-Leone, JA (1970). A mathematical model for the transition rule in Piaget's developmental stage. *Acta Psychologica*, *32*, 301–345.

Pashler, H (1998). *The Psychology of Attention*. Cambridge, MA: MIT Press.

Rohrer, D and Pashler, HE (2003). Concurrent task effect on memory retrieval, *Psychonomic Bulletin and Review*, *10*(1), 96–103.

Rohrer, D, Pashler, H and Etchegaray, J (1998). When two memories can and cannot be retrieved concurrently. *Memory and Cognition*, *26*(4), 731–739.

Rosen, VM and Engle, RW (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211–227.

Schneider, W and Detweiler, M (1987). A connectionist/control architecture for working memory. In GH Bower (ed), *The Psychology of Learning and Motivation*, *Vol. 21* (pp. 54–119). New York: Academic Press.

Siegel, LS (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioral Development*, *17*(1), 109–124.

Towse, JN and Hitch, GJ (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *The Quarterly Journal of Experimental Psychology*, *48A*, 108–124.

Towse, JN, Hitch, GJ and Hutton, U (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language*, *39*, 195–217.

Towse, JN, Hitch, GJ and Hutton, U (2000). On the interpretation of working memory spans in adults. *Memory and Cognition*, *28*, 341–348.

Towse, JN, Hitch, GJ and Hutton, U (2002). On the nature of the relationship between processing activity and item retention in children. *Journal of Experimental Child Psychology*, *82*(2), 156–184.

Turner, ML and Engle, RW (1989). Is working memory task dependent? *Journal of Memory and Language*, *28*, 127–154.

Winkelman, JH and Schmidt, J (1974). Associative confusions in mental arithmetic. *Journal of Experimental Psychology*, *102*(4), 734–736.

Zbrodoff, NJ and Logan, GD (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General*, *115*, 18–130.