

Promoting the experimental dialogue between working memory and chunking: Behavioral data and simulation

Sophie Portrat^{1,2,4} · Alessandro Guida³ · Thierry Phénix^{1,2} · Benoît Lemaire^{1,2}

Published online: 23 November 2015
© Psychonomic Society, Inc. 2015

Abstract Working memory (WM) is a cognitive system allowing short-term maintenance and processing of information. Maintaining information in WM consists, classically, in rehearsing or refreshing it. Chunking could also be considered as a maintenance mechanism. However, in the literature, it is more often used to explain performance than explicitly investigated within WM paradigms. Hence, the aim of the present paper was (1) to strengthen the experimental dialogue between WM and chunking, by studying the effect of acronyms in a computer-paced complex span task paradigm and (2) to formalize explicitly this dialogue within a computational model. Young adults performed a WM complex span task in which they had to maintain series of 7 letters for further recall while performing a concurrent location judgment task. The series to be remembered were either random strings of letters or strings containing a 3-letter acronym that appeared in position 1, 3, or 5 in the series. Together, the data and simulations provide a better understanding of the maintenance mechanisms taking place in WM and its interplay with long-term memory. Indeed, the behavioral WM performance lends evidence to the functional characteristics of chunking that seems to be, especially in a WM complex span task, an attentional time-based mechanism that certainly enhances WM performance but also competes with other processes at hand in WM. Computational

simulations support and delineate such a conception by showing that searching for a chunk in long-term memory involves attentionally demanding subprocesses that essentially take place during the encoding phases of the task.

Keywords Working memory · Chunk · Computational model · Time-based resource sharing model · TBRS*

Working memory (WM) refers to a cognitive system devoted to the simultaneous maintenance and processing of information (Baddeley & Hitch, 1974). The capacity and performance of this system is crucial because it is strongly linked to high-level cognitive activities such as reasoning, reading comprehension, and problem solving (Kane, Conway, Hambrick, & Engle, 2006; Kyllonen & Christal, 1990).

Since Daneman and Carpenter's (1980) seminal attempt to capture WM limitations with the reading span task, several tasks have been developed. Among them, the computer-paced WM complex span task (WM-CST) has been demonstrated to be highly linked to high-level cognition (Barrouillet, Lépine, & Camos, 2008) and very fruitful, especially in uncovering crucial latent dimensions such as the pace at which the concurrent activity has to be performed (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007; Barrouillet, Portrat & Camos, 2011). The paradigm consists of a dual task alternating between memory and processing phases, the duration of which is controlled. Participants have to memorize series of items (e.g., letters, digits, locations) for further immediate recall while processing other stimuli (e.g., words, digits, squares) interleaved between memory items. The nature of the concurrent processing task can be varied (e.g., reading, parity judgment, spatial judgment). Hence, in the WM-CST, which is the

✉ Sophie Portrat
Sophie.Portrat@upmf-grenoble.fr

¹ LPNC, University Grenoble Alpes, F-38040 Grenoble, France

² CNRS, LPNC UMR5105, F-38040 Grenoble, France

³ CRPCC, Université Rennes 2, Rennes, France

⁴ Laboratoire de Psychologie et NeuroCognition, Bat. Sciences de l'Homme et Mathématiques, Université Pierre Mendès-France, BP 47, 38040 Grenoble Cedex 9, France

base of the present study, distractors are interspersed between memoranda.

In parallel, the exploration of limitations in the human capacity to store information was marked by the famous magical number proposed by Miller (1956). He suggested that the short-term capacity is limited to 7 ± 2 items. But he also observed that this ability is not large enough to be efficient in daily life. For example, imagine that you have to memorize a phone number during a conversation. It is a series of 10 digits in France (e.g., 0165120417). Miller proposed that you are able to group sets of familiar items and to organize them into units called chunks. Hence, the 10 digits can be grouped in pairs: 01 65 12 04 17. But larger grouping can be realized, for example by identifying the birth date of your child. The result would be only four chunks to maintain: 01 65 1204 17. For Miller, this recoding mechanism is “*an extremely powerful weapon for increasing the amount of information that we can deal with*” (Miller, 1956, p. 95). Even if the WM storage limitation and its value were discussed in recent decades, a WM capacity of about three or four chunks seems now to be widely accepted (Chen & Cowan, 2005; Cowan, 2001; Cowan, Chen, & Rouder, 2004; Gobet & Clarkson, 2004; Mathy & Feldman, 2012).

Chunking was initially proposed as an adaptive mechanism based on expertise by researchers trying to explain the extreme capacities of chess players (Chase & Simon, 1973; De Groot, 1978; Gobet et al., 2001). Nowadays, the concept has largely spread out the domain of expertise, and two types of chunking can be distinguished (e.g., Gobet et al., 2001): implicit chunking, especially in the expertise tradition (e.g., Chase & Simon, 1973) and goal-directed chunking (e.g., Miller, 1956). It is this latter approach that has mainly been adopted by the WM community (e.g., Cowan, 2001, 2005a) and in the present paper.

Even if chunking is well known, when it comes to WM paradigms, chunking is more often used to explain data than explicitly investigated. When studying the effects of chunks in immediate memory,¹ paradigms with sequential or simultaneous presentation of memoranda are preferred (e.g., Bor & Owen, 2007; Chen & Cowan, 2005; Cowan et al., 2004; Gobet & Clarkson, 2004; Mathy & Feldman, 2012; but, see Naveh-Benjamin, Cowan, Kilb, & Chen, 2007). The consequence is that WM and chunking have been living parallel experimental lives. Only a few studies have explicitly examined what happens when elements of a chunk are not directly accessible together in the environment because of memory distractors interspersed in between. Is the chunking mechanism efficient enough to enhance memory capacity even when input conditions are not favorable to grouping, as in a WM-CST paradigm?

¹ Immediate memory is used as an umbrella term for short-term memory and working memory.

The aim of this study is (1) to strengthen the experimental dialogue between WM and chunking by studying the effect of chunking in a computer-paced WM-CST paradigm and (2) to formalize explicitly this dialogue within a computational model. To do so, we postulated that, in conditions where memory items are scattered due to interleaved processing activities, chunking would be resource demanding and its efficiency would then depend on time constraints. We thus propose a functional conception of the time course of such a chunking mechanism by supplementing the time-based resource sharing model (TBRS; Barrouillet et al., 2011) with an additional chunking mechanism that, while being considered as a memory enhancer, would nevertheless compete with the refreshing process because it also consumes attention. In order to better formalize this conception, we append a chunking module to TBRS*, a computational implementation of TBRS (Oberauer & Lewandowsky, 2011).

WM maintenance mechanisms contributing to recall performance

Even if chunking can be considered, as in the present paper, to be a mechanism that would participate in the recall performance observed in WM tasks (e.g., Cowan & Chen, 2009; Stuart & Hulme, 2009), classically, two other verbal maintenance mechanisms are put forward: phonological rehearsal and attentional refreshing. Evidence for the existence of these two systems of maintenance in verbal WM has recently been accumulated (Camos, Lagner, & Barrouillet, 2009; Camos, Mora, & Barrouillet, 2013; Camos, Mora, & Oberauer, 2011; Mora, & Camos, 2013). The phonological or articulatory rehearsal was initially described by Baddeley (1986) and consists of the vocal or subvocal repetition of the items to be maintained within a loop. In accordance with Baddeley’s idea of a WM multicomponent system, the phonological rehearsal is a domain-specific mechanism peculiar to verbal material. The second maintenance mechanism at hand in WM is the attentional refreshing mechanism described as a covert retrieval (Cowan, Saults, & Elliott, 2002; Hudjetz & Oberauer, 2007), and viewed as a general attention-based system in several theories (Barrouillet et al., 2011; Camos et al. 2011; Cowan, 1999, 2005a; Johnson, 1992; Johnson et al., 2005; Raye, Johnson, Mitchell, Greene, & Johnson, 2007). Thanks to such a mechanism, memory traces are maintained in an active state through attention. Focusing briefly on an item enhances its level of activation and thus counteracts memory loss. It has been suggested that these two classic maintenance mechanisms, operating independently on the maintenance of verbal information, affect different features of the memoranda (Camos et al., 2009; Camos & Portrat, 2015).

As a memory enhancer, chunking could be most likened to refreshing than to rehearsal because of its link to attention

(Cowan & Chen, 2009) and because of the experimental dissociation found between chunk representations and phonological representations (Chen & Cowan, 2005; Cowan et al., 2004). According to Cowan (1999, 2001, 2005a, 2005b), the focus of attention is used to remember information, and the process by which this occurs involves new long-term memory (LTM) formation. New associations are formed between elements concurrently held in the focus of attention resulting in new LTM information that is formed during short-term memory tasks: a chunk. The information that benefits from attentional refreshing may be in a more deeply analyzed form (Cowan & Chen, 2009) and more persistent (Camos & Portrat, 2015) than information that is activated without the involvement of the focus of attention, that is, by articulatory rehearsal. Hence, associations or links between elements, that is, chunks, are stored as a specific function of the focus of attention (Cowan, 1999, 2001, 2005a, 2005b) and the focus of attention would be limited to a few chunks at a time (usually three to five chunks in adults; see Cowan, 2001).

A consequence of the attentional conception of chunks and chunking is that to obtain a *chunk effect* that would enhance memory capacity, the paradigm would have to fulfill an important condition: allowing the different constitutive elements of a chunk to be under attentional focusing at the same time. Hence, classically, items that constitute a chunk are either simultaneously presented or in immediate succession (e.g., Bor & Owen, 2007; Cowan et al., 2004; Chen & Cowan, 2005; Gobet & Clarkson, 2004; Mathy & Feldman, 2012). However, as we have seen earlier, a well-established WM paradigm that has shown its ability to predict high-level cognition is precisely one in which processing phases are interleaved between successive memoranda. Hence, we will present, in the next section, several assumptions that would make it possible to study the time course of a chunking mechanism taking place in a WM-CST paradigm and will propose their implementation in a computational model.

The time course of chunking in a WM-CST

The time-based resource-sharing model (TBRS; Barrouillet et al., 2004; Barrouillet et al., 2007; Barrouillet et al., 2011) and its computational implementation (TBRS*; Oberauer & Lewandowsky, 2011) constitute the background of our investigations because they are excellent bases for studying the time course of mechanisms at hand in WM. TBRS describes the functioning of WM when participants are performing a WM-CST. This paradigm consists in a dual task in which participants memorize items while concurrently performing a distracting activity. Participants are thus presented with memory items for further recall, each item being followed by a concurrent processing task which could be varied (e.g., reading, spatial judgment, parity judgment). The bottom part of

Fig. 1 depicts a fictional situation in which participants have to maintain four letters (K, P, D, and F) and concurrently to judge the position (upper or lower part of the screen) of two successive squares appearing after each letter. Basically, TBRS supposes that after being encoded, memory traces of the items to be maintained are kept vivid through attentional focusing (*refreshing*) but fade away as soon as they leave the focus of attention. This is exactly the case when a concurrent processing activity occurs because of a central bottleneck (Pashler, 1998) that allows only one elementary cognitive step to take place at a time. Hence, during intervening activities, memory traces suffer from a *time-related decay*, but they can be refreshed through attentional focusing during short pauses that would be freed during the processing phase. Fig. 1 shows the time course of these processes through the evolution of the level of activation of the four letters to be maintained in WM.

According to this functional conception, within a WM-CST, WM performance depends on the cognitive load of the processing task that corresponds to the proportion of time during which controlled attention is captured by the concurrent activity. In other words, the more the distracting task captures attention, the less the memory traces can be refreshed and hence the poorer the recall performances. Several experimental pieces of evidence have been collected in favor of this theoretical model of WM via behavioral studies in adults and children (e.g., Barrouillet et al., 2007; Barrouillet et al., 2011; Gaillard, Barrouillet, Jarrold, & Camos, 2011; Portrat, Camos, & Barrouillet, 2009; Vergauwe, Barrouillet & Camos, 2010).

Within such a model, *three assumptions* could be made about the time course of the chunking mechanism. First of all, *chunking takes place within the focus of attention*. Hence, because of the attentional bottleneck, chunking would occur only when no other attention-demanding activity is running. In this sense, chunking would suffer from exactly the same constraints as the attentional refreshing described above and would thus take place only during short pauses that would be freed. Therefore, the two attentional maintenance mechanisms, refreshing and chunking, would be in competition with each other. Second, *all the constitutive elements of a chunk would have to be simultaneously present within the focus of attention* (Cowan & Chen, 2009). As a consequence, within a WM-CST, at the time of encoding a given memory item, the preceding ones would have to be reactivated for the sake of potential chunking. It follows that the focus of attention has to be able to handle more than one item at the same time. This point of view is not at odds with the TBRS conception of WM functioning. Indeed, according to Pashler (1998), TBRS postulates that a central bottleneck allows only one controlled elementary process to take place at a time (e.g., Barrouillet et al., 2007). However, this does not mean that each single elementary process cannot handle several pieces of information at the same time. This conception can also be compatible with the concentric model of WM proposed by Oberauer

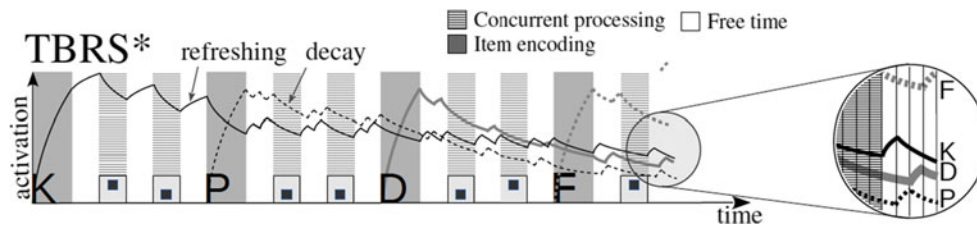


Fig. 1 This diagram illustrates the evolution of the item/position activation values in the TBRS* model. For the sake of clarity, several simplifications have been made: There are only two distracting episodes after each of the items to be maintained (instead of four in the present study), there are no retrieving errors (while the model actually sometimes mistakenly retrieves wrong items), and the different durations do not correspond to the real ones (e.g., the attentional capture imposed by the concurrent processing task during which the items decay is actually shorter than the duration of the presentation of the distracting stimulus). The first item to be maintained is K. The light gray area represents the encoding step during which the activation of K is increased. The next

white area represents the time devoted to refreshing items. Since there is only one item (K) encoded so far in WM, the activation value continues to increase until a distracting episode occurs (horizontal hatching area): activation decays. The task then alternates between free time (refreshing) and distractors (decay) until a new item, P, occurs. During the free time following the encoding of P, as well as during the free time following a concurrent processing episode, the two items (P and K) are refreshed in turn: When one is refreshed, the other one decays. And so on until the end of the series. The magnifying circle at the right side of the diagram emphasizes the refreshing of the four items, one after the other

(2002). Because even if in this model, only one item can be selected for processing in the focus of attention, a region of direct access supplements the focus of attention and “holds a limited number of chunks available to be used in ongoing cognitive processes” (Oberauer, 2002, p. 412). In such a theory, chunking would take place in that region of direct access.

The third and final assumption is related to the time constraints that are inherent to the WM-CST. As is the case in several models including a chunking mechanism (e.g., Jones, Gobet, & Pine, 2007; Richman, Staszewski, & Simon, 1995; Waters & Gobet, 2008), we postulate that the chunking mechanism would involve several sub-processes like reactivation of previous memory items, comparison with LTM knowledge, chaining and recoding, each of them being time-consuming.

TBRS*, the computational implementation of TBRS, is fed with behavioral data collected from a WM-CST (Oberauer & Lewandowsky, 2011; Portrat & Lemaire, 2015). Basically, this model is made up of a data structure and four basic operations. The data structure of the model is a two-layer connectionist network in which one layer is dedicated to the memory items and the other layer represents the position of each item in the series. The two layers are fully interconnected. The four operations enable the model to (i) encode an item at a given position, (ii) process distractors, (iii) refresh memory traces, and finally (iv) recall items. Our goal was to supplement TBRS* with a chunking mechanism in order to replicate the behavioral memory performance of participants in a WM-CST favoring chunking. This goal had to be reached with as few modifications as possible compared with the original model. To this end, the original data structure remained the same without creating any new units in the item layer that would have represented chunks. However, the four basic functions were adapted to allow the grouping of isolated items into chunks when necessary.

In the original TBRS* model (Oberauer & Lewandowsky, 2011), each position is coded by a subset of position units, so that two adjacent positions share a proportion of P units. Encoding (i) is modeled as a process of connecting positions with items, by Hebbian learning. The strength of the increase of a connection weight (w) depends on a strength parameter (η) and it is bound by an asymptote L .

$$\Delta w = (L - w)\eta. \tag{1}$$

The strength η depends on the time (t) devoted to encoding as well as a stochastic parameter (r) modeling human variability:

$$\eta = 1 - e^{-rt}, \tag{2}$$

with $r = N(R, \sigma^2)$. Usually, $R = 6$, $\sigma = 1$, and $t = 500$ ms.

For instance, if the sequence to be memorized is KPDP (as in Fig. 1), K is first encoded which results in strengthening the links between item K and the nodes coding for Position 1.

When participants have to process distractors (ii), attention is captured by the other task, and those link values (w) decrease according to an exponential function (decay):

$$w(t) = w_0 e^{-Dt}. \tag{3}$$

D is usually set at $1/2$.

When attention is redirected toward the memory task, a refreshing process (iii) takes place and leads to an increase of the w values. Each atomic refreshing follows Eq. 2 but with a much shorter time Tr (usually 80 ms), compared with the initial encoding ($t = 500$ ms). Following a classic loop conception, all items are successively considered, starting with the first one and this process cycles until a new activity requires attention. To pursue our example, when P is encoded, activation values between the node representing P and the

nodes representing Position 2 are strengthened (while in the meantime, the activation values of K are decreased). If there is time for refreshing, it is done alternately between P and K. In a recent modified version of TBRS*, we demonstrated that this classic loop conception of refreshing is not suitable to account for serial position curves observed on adults' recall (Portrat & Lemaire, 2015). Hence, we proposed an alternative to the purely sequential refreshing mechanism handling one item at a time during an average duration of T_r ms: up to four items can be refreshed at the same time. The duration of each refreshing step is still set at T_r ms (80 ms) but the strength of the links between the items and their position units (parameter η in Eq. 1) is therefore divided by the number of items simultaneously handled, since the attention involved in such controlled processes is a limited resource (e.g., Barrouillet et al. 2007; Engle, Kane, & Tuholski, 1999).

When it is time to *recall* items (iv), the one most connected with the position at hand is recalled. To minimize the accessibility of that item later in the same series, all its connections to the position layer are decreased by Hebbian antilearning.

In this model, apart from the successive positions that shared common units, there is no consideration for associations between memory items. Hence, chunking cannot be directly simulated in such a framework. Yet, in memory tasks as well as in, particular, WM-CST's, memory items are obviously linked to each other, over and above their specific position. It is very likely that, even if memory items are scattered, items themselves, rather than their position in the list, are more or less linked together to form meaningful groups (i.e., chunks). Hereafter, we will propose some modifications to the original mechanisms that would make it possible to simulate chunking.

Implementation of the chunking process in a complex span task: TBRS*C

Chunking is a process that relies on a storage system, external to WM, which maintains known sequences of items (Gobet et al., 2001). There is a long history of research aiming at modeling the relationship between chunking and this LTM storage system. This line of research gave birth to computational models such as Soar (Laird, Rosenbloom, & Newell, 1986), ACT-R (Anderson et al. 2004) or CHREST (Gobet et al., 2001). These models consider that chunking is a general learning mechanism of human cognition that tends to group and memorize, in LTM, units that occurred frequently or actions that happened to be successful. Modeling the storage and retrieval into LTM is therefore crucial for those models. Our goal is somehow different because our focus is on WM. Although we obviously need to implement an LTM system, the storage and retrieval into that structure were not investigated in the present study. Therefore, our model describes an

LTM as simple as possible consisting of a list of objects that correspond to groups of letters such as PDF or JPG. When simulating a specific experiment, this list is initialized with all sequences of the experimental material that participants are likely to recognize as chunks. Figure 2 depicts the time course of processes taking place in TBRS*C (for TBRS* with chunks) when submitted to a WM-CST in which there are acronyms.

Right after encoding a new item, TBRS*C searches to see whether this item, when concatenated with the previous ones, would form a chunk. In the experiment presented below, chunks could only be three letters long, and there could be only one chunk in each of the series to be maintained. Therefore, the current version of TBRS*C initiates this search only after three items have been encoded and if no chunks have been identified so far. In these cases, TBRS*C first retrieves items in positions N-2 and N-1. These retrievals are made, as usual, by activating the corresponding position units and selecting the most activated item. These reactivations of previously encoded memory items correspond to our above assumptions according to which the constitutive items of a given chunk should be readily accessible to allow the chunking process. This implementation implies, as proposed by Cowan (2005a, b) and as we already modeled (Portrat & Lemaire, 2015), that the focus of attention can handle more than one item at a time. Then, the LTM knowledge component is queried with this sequence of three items. As mentioned previously, this process takes time and its duration is controlled by a parameter called *chunk Search Duration* (*cSD*). LTM chunk retrieval does not always succeed. Therefore, we defined a parameter called *Probability of Chunk Retrieval* (*PCR*) which controls the likelihood of a sequence of items existing in LTM being recognized as a chunk. For example, if *PCR* is set to .33 and "JPG" has been defined as an LTM chunk, the model would recognize the sequence J-P-G as a chunk in only one case out of three. It is worth noting that these two free parameters (*cSD* and *PCR*) are independent of each other with their own *raison d'être*. On the one hand, *cSD* is a way by which we modeled the time constraints and the resource consuming assumptions inherent to the chunking mechanism described above. On the other hand, *PCR* is necessary to model the fact that chunks are not always retrieved. One might have thought that *cSD* and *PCR* values were inescapably dependent on each other, since a longer time spent searching for a chunk would lead to an increased probability of retrieving it. However, we opted for a separate *PCR* parameter because it is also possible to spend a long time searching for a chunk without necessarily finding it when, for example, the acronym is rare. Even if this factor has not been investigated in the present study, the current model is already able to tackle this issue.

When the LTM knowledge component returns that the sequence of items is known, that chunk is now considered as a

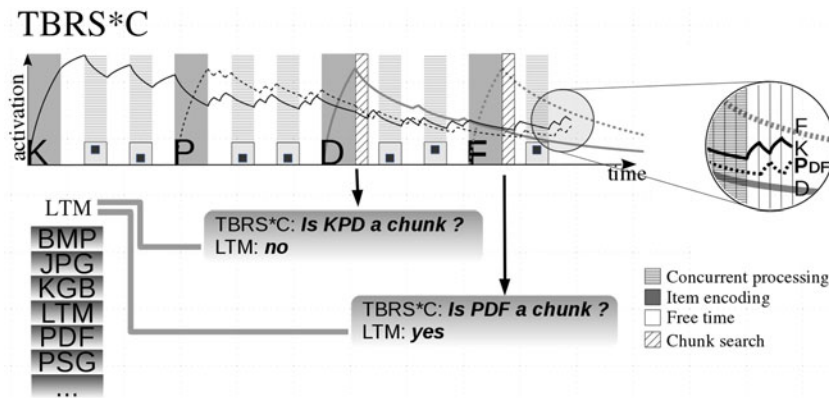


Fig. 2 This figure illustrates how TBRS*C works, in comparison with TBRS*. The evolutions of activation values are identical until three items have been encoded. After D has been encoded, TBRS*C spends time checking whether the preceding three items (K, P, and D) form a chunk. LTM is thus queried and the answer is no because this three-letter series is not stored in the LTM component. During that searching time (corresponding to the *cSD* parameter), the activation values of the three letters already encoded decay. Indeed, no item can be refreshed during the chunk search. In this particular (and fictional) situation, the remaining

short free time after searching for a chunk allows only the first two items to be refreshed: D still decays and cannot be refreshed. Right after encoding F, a new query is made and the answer is positive: PDF belongs to LTM. The three items are then chained together. Only two units now have to be maintained: K and PDF. The magnifying circle at the right side of the diagram illustrates that point: D and F are no longer refreshed as single items and they decay. Only K and P (chained with D and F) are considered and refreshed in turn

single item and reencoded in the position of its first item. For instance, consider the sequence KPDPF as in Fig. 2. K, P and D have already been encoded, and F is now presented. Suppose also that after encoding F, P has been retrieved in Position 2 as well as D in Position 3. The LTM knowledge component is queried with “PDF,” which returns true, and the chunk “PDF” is reencoded in Position 2. The three items constituting the chunk are then chained together within the item layer in order to be retrieved as a single object.

To counteract memory loss, a refreshing mechanism occurs when attention is not captured by concurrent processing activities. Each position is activated in turn, starting with the first one, exactly like refreshing in TBRS*. If what is retrieved is a chunk, it is refreshed exactly as a single item. This feature of TBRS*C is akin to the one proposed in EPAM IV, in which the short-term memory component is limited by the need to rehearse chunks before they fade from memory (Richman et al., 1995); in CHREST, in which representations linked to chunks decay if not refreshed (Waters & Gobet, 2008); as well as in EPAM-VOC, in which the information, such as a chunk, held in phonological WM is subject to decay within 2 seconds (Jones et al., 2007). It is worth noting that a previously encoded chunk is not always retrieved at the time of refreshing because of the random noise or the overlapping of adjacent positions, two phenomena that generate possible memory loss. After a chunk is refreshed, the next position under consideration is the current position + 3 (or the Position 1 if the end of the sequence is reached). For instance, if the current sequence is KPDPFV, and Position 2 is activated, and “PDF” is retrieved and refreshed, the next position considered for refreshing is Position 5, which would probably retrieve V. Since the attentional focus size is four, both K, PDF, and V are within

the focus (together they all represent only three units) and refreshed together. It is worth noting that the links between Position 3 and D, as well as between Position 4 and F, are no longer reactivated and would decay over time. Like other items, chunks decay over time with the same decay rate as single items.

We also implemented the fact that participants would search, too, for chunks during free time in between successive processing activities. For instance, if we consider that four items have been encoded so far, free time would be occupied as such: retrieving items at Positions 1, 2, 3, and 4, refreshing them, querying the LTM knowledge component to check whether the two 3-item sequences (in Positions 1, 2, and 3 and Positions 2, 3, and 4) form a chunk, and so forth. We therefore defined a third parameter, *chunk Search Duration in Refresh (cSDR)*, controlling the duration of that search process during the processing phase.

These subprocesses that supplemented the original TBRS* to manage with the chunking mechanism are in line with our theoretical assumptions. According to these assumptions, chunking, when inserted in a WM paradigm, is a process that takes time and competes with distracting processing activities and item refreshing. To sum up, chunking is considered as a maintenance mechanism that enhances memory performance. It consists in grouping together several pieces of information in order to constitute a unique significant unit. We propose that chunking is composed of subprocesses: a *recognition* process that requires a *reactivation* of previously encoded items associated with a *maintenance* of the detected chunk in memory. We propose that within a WM-CST, with scattered presentation of memoranda, chunking is not an automatic mechanism but is instead constituted of subprocesses that rely

on controlled attention. This activity should therefore compete for this limited resource with the other demanding activities at hand in WM-CST: maintaining memory items and processing concurrent information. Moreover, its efficiency should depend on the time available for maintenance mechanisms within the task.

In order to tap the chunking mechanism, acronyms were inserted within strings of letters to be remembered. First, when acronyms are present in the memoranda, memory performance should be better than when they are not. In fact, instead of refreshing three dissociated elements, only one element (the chunk) needs attention. This simple effect is hereafter called the chunking effect. Second, we hypothesized that the position of the acronym in the series should have an effect on recall performance. Indeed, we postulated that the earlier the acronym was presented within the serial positions, the more beneficial the chunking should be because there are more opportunities to find a chunk and to refresh one element (the chunk) instead of three. Finally, the classic pace effect that reveals time-based switching between processing and maintenance activities should be observed in the conditions with no acronyms as well as in the conditions with acronyms, since the chunking process is, especially in the present paradigm, an attention demanding processing that would take place only when attention is not captured by another activity (processing distractors or refreshing individual memoranda).

To test these hypotheses, we proposed to adult participants a computer-paced WM-CST in which they had to remember series of seven letters while performing a concurrent interleaved location judgment task consisting in evaluating the relative position of four black squares that appeared successively on screen after each letter (Portrat et al., 2008). In some cases, the memoranda contained three-letter acronyms while in other cases acronyms were deliberately avoided. Moreover, to test the time-based hypothesis, we manipulated the pace at which the concurrent processing task had to be performed (fast or slow pace) as well as the position of the acronym within the series (beginning, middle, end).

Experiment

Method

Participants Thirty-nine undergraduate psychology students at the University of Grenoble (23 women) received partial course credit for participating. Their mean age was 21.6 years ($SD = 2.2$), and 89% were right-handed. Seven participants were excluded from the following analyses (more details are given in the “results” section).

Materials We created specific material for the memory task that included three-letter French acronyms. First of all, 100

students, who were not engaged in the rest of this study, were asked to evaluate 68 acronyms according to this grid: *unknown*, *known but not able to define*, and *known and able to define*. From this pretest, the 24 acronyms that received the most *known and able to define* responses were selected to constitute the acronyms of the WM-CST. The chunk recognition rate was from 100% (PSG for Paris Saint Germain) to 82% (BSR for Brevet de Sécurité Routière) with an average at 96%.

For the WM-CST, participants had to perform three trials for each of the eight experimental conditions resulting from the orthogonal manipulation of the two variables in a within-subject design: the Pace factor (slow and fast) and the Acronym factor, according to which a three-letter acronym was either absent, in the first, third, or fifth position within the memoranda. The 24 series of seven consonants were obtained by adding four consonants to each acronym with the following restrictions: no W (which is trisyllabic in French), no repetitions, no other acronyms, and no alphabetically ordered strings. Moreover, each consonant used was equally distributed across all conditions and appeared equally across all series. Eight counterbalanced sets of 24 series were thus created. The condition with no acronym was constructed by changing the order of the three letters of a given acronym and spreading these letters through the series (see Fig. 3).

The processing task consisted in a location judgment in which participants had to decide whether a displayed square was located in the upper or lower part of the screen. The stimuli of this processing task were the same as those used in Barrouillet et al. (2007, Experiment 2, distant condition) and Portrat et al. (2008). They consisted in black squares (side = 18 mm subtending 2° in visual angle, for a participant who was seated about 60 cm from the computer screen) centered on one of two possible locations, either in the upper or the lower part of the screen. The two locations were 68 mm apart (6.5° in visual angle). Squares were randomly displayed in both locations with the same frequency.

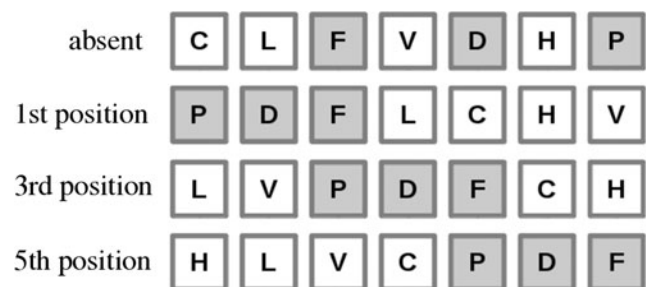


Fig. 3 Example of series to be maintained across the four conditions of the Acronym factor (absent, present in the first, third or fifth position). Each type of series was used in the two Pace conditions and counterbalanced across participants

Procedure This experiment was administered individually on a computer using the PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were presented with the 24 series of seven consonants to be remembered. Each consonant was followed by a series of 4 squares successively displayed on the screen. Each series began with a first screen indicating the Pace condition during 1,500 ms (*rapide* [fast] or *lent* [slow]). After a blank screen of 500 ms, the first letter was displayed on screen for 1,500 ms, followed by a 500 ms delay. After this postletter delay, the four successive squares were displayed on screen for 520 ms followed by a 660 ms or a 260 ms postsquare delay for the slow and the fast Pace condition, respectively. Then, the next consonant appeared for 1,500 ms and so on up to the end of the series. At this moment, the word *rappel* [recall] appeared on screen, and the participants were asked to recall letters in forward order by typing them on the computer keyboard. Then, to ensure that a goal-directed chunking mechanism had effectively taken place, participants had to say whether they had found an acronym in the series, and to recall it if such were the case. As participants are asked whether they found acronyms, they are more likely to actively and deliberately search for them. All the recall procedure was self-paced, participants pressing the space bar to go forward. Besides the letters and the chunks recalled, reaction times and accuracy during the location task were recorded.

Participants were asked to read aloud and to memorize each letter when it appeared and to judge the location of each square as quickly as possible without sacrificing accuracy, by pressing either the left-handed key “Q” or the right-handed key “M” of an “azerty” keyboard for the lower and the upper locations respectively.

The experimental session was preceded by familiarization phases. Participants familiarized themselves with the location task on 64 stimuli at slow pace first and then at fast pace. In these two training phases, participants heard a beep if they made a mistake or if they were too slow in responding (i.e., longer than the sum of the presentation duration and the postsquare delay). If they did not reach 90% correct responses, they had to perform again the same series of 64 squares with a maximum of three training phases. The last training phase aimed at familiarizing participants with the WM-CST and consisted of two trials: a slow pace / no acronym and a fast pace / acronym in third position.

Results

All the participants took part in the experimental phase, but seven of them were excluded from the following analyses: three participants because of technical problems and four others because they did not perform the concurrent processing task properly with a mean score above 80% in judging the location of squares. The remaining 32 participants performed

pretty well the concurrent task with a mean score of 96% ($SD = 3\%$) that did not vary as a function of the experimental conditions, $F(1,93) < 1$ and $F(3,93) < 1$ for the Pace and the Acronym factors respectively.

Correct recall of letters

We performed an ANOVA on the number of letters recalled in the correct order (out of seven letters) with the Pace conditions (fast and slow) and the Acronym conditions (absent, first, third, and fifth positions) as within-subject factors (see Fig. 4).

Overall, the two variables at hand had an effect on recall performances with a significant Pace effect, $F(1, 31) = 10.71$, $MSe = .841$, $p = .002$, $\eta^2 = .26$, as well as a significant Acronym effect, $F(3, 93) = 11.94$, $MSe = 1.088$, $p = .000001$, $\eta^2 = .27$, but these effects did not interact, $F(3, 93) = 1.31$, $MSe = 1.185$, $p = .276$, $\eta^2 = .04$. Furthermore, we performed several comparisons to decompose the main effect of the Acronym variable.

Concerning the Acronym dimension, the *chunk effect* comparing the condition with no acronym to the three acronym conditions in a planned comparison was significant, $F(1, 31) = 16.53$, $MSe = 1.29$, $p = .0003$, $\eta^2 = .35$. Moreover, the beneficial effect of the presence of acronyms with the memoranda varied as a function of its position in the series. Indeed, Position 1 was more favorable than Positions 3 and 5, $F(1, 31) = 16.08$, $MSe = 0.985$, $p = .0003$, $\eta^2 = .34$. Positions 3 and 5 did not lead to different recall performances, $F(1, 31) = 1.14$, $MSe = 0.987$, $p = .294$, $\eta^2 = .03$. Participants recalled on average 5.36 letters out of seven when acronyms were in Position 1 while they recalled only 4.85 and 4.66 when acronyms were in Positions 3 and 5, respectively.

In order to analyze the results more deeply, we separated two dimensions that were entangled in the previous analysis, the recall of isolated items and the recall of acronyms.

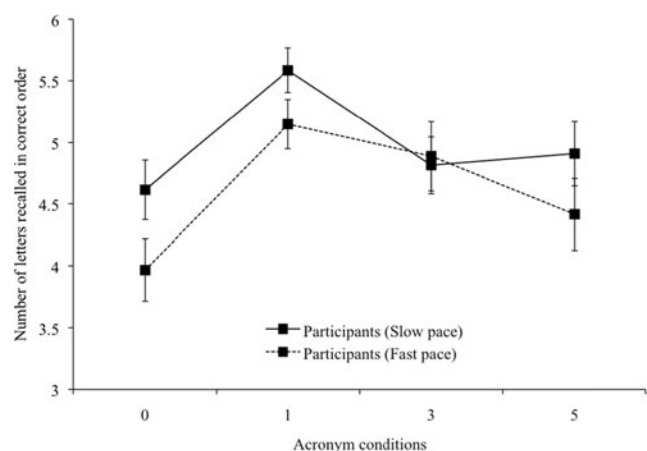


Fig. 4 Mean number of letters recalled in correct order out of seven (error bars are standard deviations) as a function of the Pace conditions (fast and slow) and the Acronym conditions (absent, first, third, and fifth positions)

Correct recall of isolated items

An ANOVA was performed on the number of isolated letters recalled in correct order (out of four letters) with the Pace conditions (fast and slow) and the Acronym positions (first, third, and fifth positions) as within-subject factors. Especially of interest here, the Pace effect when tested only on the isolated letters was still significant, $F(1, 31) = 4.60$, $MSe = .423$, $p = .04$, $\eta^2 = .13$. Concerning the rest of the analysis, the Acronym factor was also significant, $F(2, 62) = 9.48$, $MSe = .549$, $p = .0003$, $\eta^2 = .23$ and no interaction was found, $F(2, 62) = 0.81$, $MSe = .883$, $p = .449$, $\eta^2 = .03$.

Correct recall of acronyms

We also performed an ANOVA on the number of acronyms correctly recalled with the Pace conditions (fast and slow) and the Acronym position (first, third, and fifth positions) as within-subject factors. The best score on this variable is three since each participant performed three trials under each condition (see Fig. 5). The position of the acronym within the series had a large effect on performance, $F(2, 62) = 42.46$, $MSe = 0.617$, $p < .0000001$, $\eta^2 = .58$. Overall, participants recalled 2.58 acronyms out of three when the acronym was in the first position, while they recalled only 1.55 and 1.41 acronyms out of three when they were in the third and the fifth positions, respectively. Planned comparisons revealed again the superiority of the first position, $F(1, 31) = 86.09$, $MSe = .601$, $p < .0000001$, $\eta^2 = .73$, and no difference between Positions 3 and 5, $F(1, 31) = 1.00$, $MSe = .6328$, $p = .3250$, $\eta^2 = .03$. However, results did not show a significant effect of Pace, $F(1, 31) = .82$, $MSe = .4086$, $p = .3734$, $\eta^2 = .03$, nor an interaction between Pace and Acronym position, $F(2, 62) = .38$, $MSe = .5099$, $p = .6868$, $\eta^2 = .012$.

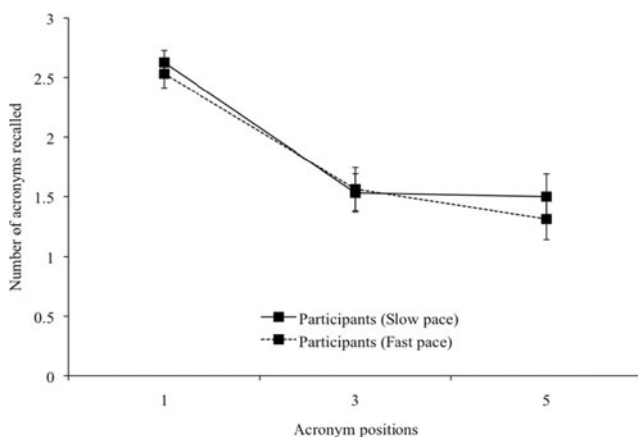


Fig. 5 Mean number of acronyms recalled out of three (error bars are standard deviations) as a function of the Pace conditions (fast and slow) and the position of the Acronym in the series to be maintained (first, third, and fifth positions)

Simulations

As mentioned earlier, we supplemented the original TBRS* with a mechanism to deal with chunks, using three new parameters: the duration of the process of searching for chunks after encoding a new item (cSD), the duration of the process of searching for chunks during the refreshing phases ($cSDR$) and the probability of recognizing a chunk in LTM (PCR).

In order to assign relevant values to the three new parameters, we performed a grid search in this three-dimensional space, using the default values of the original TBRS* model parameters. Parameter ranges were the following; $cSD \in [0 \text{ ms}, 200 \text{ ms}, 400 \text{ ms}, \dots, 2,000 \text{ ms}]$, $cSDR \in [0 \text{ ms}, 50 \text{ ms}, 100 \text{ ms}]$, $PCR \in [1, .9, .7, .5, .3, .1]$. We ran 10,000 simulations for each combination of parameters and computed the likelihood of each one given the experimental data. Computing likelihoods requires defining probability distributions. Following Lewandowsky and Farrell (2011), we computed for each of the eight conditions (fast or slow pace, chunk absent or in Position 1, 3, or 5) and each of the seven item positions (from 1 to 7), a probability distribution across the number of correctly recalled items at the current position over the three trials, using a binomial data model on the average percentage of correct responses for the model under consideration. For instance, the model $M(cSD = 600 \text{ ms}, cSDR = 0 \text{ ms}, PCR = .5)$ has an average percentage of correct responses in Position 4 of 40.35% in a given condition (fast pace, chunk at Position 3). Figure 6 presents the probability distribution for that model having a mean performance of 40.35% of letters recalled in that position in this condition. According to the binomial distribution, the probability of recalling only one item correct out of the three trials for that particular model is .43, which is the most probable outcome (cf. Fig. 6). However, recalling all the three items is much less likely to happen ($p = .066$).²

These probability distributions obtained for each condition in each position had been used together with the recall performance of each participant to compute the likelihood of each model. For instance, if participant #1 has correctly recalled the 3 items out of the 3 trials in position 4 in the condition described previously, that likelihood would be .066, which would not be a positive contribution to that model. On the contrary, if participant #2 has correctly recalled only one item out of the 3 trials, the specific likelihood would be .431 which would be a much more positive contribution to the overall likelihood. The likelihood of a given model corresponds to the product of the individual likelihoods. The higher that value, the better the fit. However, as is usually done, log-

² As a reminder, the probability of recalling exactly k items out of n , given the probability p of correctly recalling a single item, is $C(n; k) \cdot p^k \cdot (1 - p)^{n - k}$.

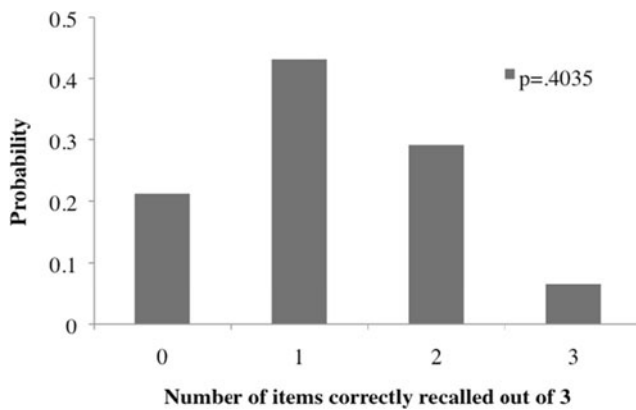


Fig. 6 Probability of correctly recalling N out of three items when the proportion of correct responses at a given position is $p = .4035$

likelihoods are considered instead of simple likelihoods and thus, the best models are the ones which have the lower negative log-likelihoods (Lewandowsky & Farrell, 2011).

Once the log-likelihoods were computed for all models (each one corresponding to a different combination of parameters), we analyzed these data. First, all models spending time to search for chunks during the refreshing phases ($cSDR > 0$), whatever the two other parameters, show log-likelihoods that are worse than all models which do not search for chunks during refreshing. In other words, chunks should be searched for right after encoding an item but not during the free time in the processing phase. The reason is that refreshing is a crucial step which has to consider all items in turn in order to increase slightly their activation value. Spending time searching for chunks therefore has a detrimental effect on the maintenance of single items and leads to performance that does not mimic that of the participants. Thus, only two parameters remain to be investigated. Figure 7 shows the negative log-likelihoods as a function of these parameters.

It turns out that there is a set of models that have low negative log-likelihoods. They lie along a valley that stretches diagonally: either a short duration of searching for a chunk with a low probability of recognizing it or a long duration of searching for a chunk with a high probability of recognizing it. The models which search for chunks for a long time without recognizing it or those which do not spend time searching but still easily find chunks do not fit well the behavioral data (the two hills on the left and right sides of Fig. 7). There are thus two kinds of best models. The first one (Model M1 hereafter) spends almost all the time between encoding an item and the first following distractor searching to see whether the current three-item-long sequence is an acronym or not. And, it has a high probability of recognizing a chunk if it exists. The benefit of the chunk therefore compensates the higher decay of the other items to be maintained that occurred during that long search process. The second kind of good model (Model M2 hereafter) does not spend much time searching for a chunk and

it does not find it very often. It could use the remaining time left free before the first distractor to refresh individual items which guarantee a reasonable performance though.

In order to select which of these two models M1 ($PCR = .8$, $cSD = 1,800$ ms, $cSDR = 0$) or M2 ($PCR = .1$, $cSD = 600$ ms, $cSDR = 0$) better reproduces human behavior, we assessed their ability to reproduce the data presented above: the seven-item recall, the recall of chunks, as well as the recall of isolated letters. We computed the root mean square error (RMSE) between models and data as well as the log-likelihood ($-\log L$), using a binomial data model as presented above, except that it is based on four data per trial for isolated letters and one datum per trial for chunks.

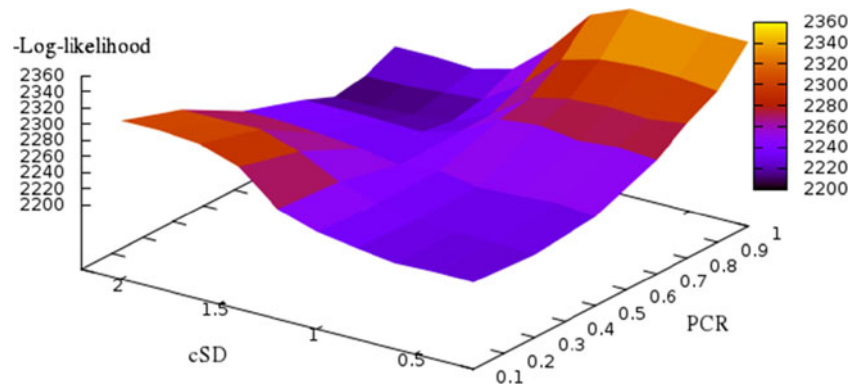
It turned out that the two models are equivalent with respect to the recall of chunks ($RMSE = .128$, $-\log L = 283.47$ for M1, $RMSE = .128$, $-\log L = 282.73$ for M2). With respect to the recall of isolated letters, Model M1 ($RMSE = .123$, $-\log L = 936.94$) appeared to be better than Model M2 ($RMSE = .123$, $-\log L = 961.22$). Finally, the measures of fit performed on the entire seven-item recall data indicated that the best fitting of human performance is also obtained by Model M1 which searches for a long time with a high probability of recognizing a chunk ($RMSE = .110$, $-\log L = 2,208.83$). Model M2 with a low probability of recognizing a chunk, fits less well with the behavioral data ($RMSE = .113$, $-\log L = 2,226.49$). The default parameters we suggest for our chunking module are therefore those of the model M1: $PCR = .8$, $cSD = 1,800$ ms, $cSDR = 0$ ms.

Moreover, the serial position curves (SPC) illustrate well the superiority of that model M1 having a high probability of recognizing a chunk (see Fig. 8). Indeed, overall, TBRSC produced nice SPCs' that replicate the classic recency and primacy effects also observed on behavioral data. Most importantly, the model M1 with $PCR = .8$ and $cSD = 1,800$ ms is able to replicate very well the specific plateaus observed in human behavior for Positions 1, 2, and 3 when the acronym appeared at the beginning of the series. Hence, as shown in the middle panel of Fig. 7, this model replicates well several aspects observed in human behavior: primacy, recency, and the facilitatory effect of acronyms in Position 1. This is not the case for the Model M2 with $PCR = .1$ and $cSD = 600$ ms (see Fig. 8, bottom panel).

Discussion

The aim of the present paper was (1) to strengthen the experimental dialogue between WM and chunking, by studying the effect of acronyms in a computer-paced WM-CST and (2) to formalize explicitly this dialogue within a computational model. Broadly speaking, our main conclusions are the following. The behavioral WM performance lends evidence to the functional characteristics of chunking, which seems to be,

Fig. 7 Negative log-likelihood as a function of the duration of the process of searching for chunks after encoding a new item (cSD) and the probability of recognizing a chunk in LTM (PCR)



especially in a WM-CST, an attentional time-based mechanism that certainly enhances WM performance but also competes with other processes at hand in WM. Computational simulations support and delineate such a conception by showing that searching for a chunk in LTM involves attentionally demanding sub-processes that essentially take place during the encoding phases of the task. However, during the processing phase, chunk maintenance appears to be roughly similar to individual item maintenance. The particular implications of these behavioral and computational results are discussed hereafter.

First of all, behavioral data confirmed that chunking enhances WM performance even in a situation in which the dispersed inputs do not favor grouping individual elements into a significant memory unit. In fact, results showed that, in a WM-CST in which each memory item is interleaved with processing phases, recall performance was better in conditions with acronyms than in a standard condition with no acronyms. The present experimental finding is thus in accordance with Cowan's theory (1999, 2001, 2005a, b) and supports the idea that chunking is a general maintenance mechanism that does not require the simultaneity of inputs.

Second, the advantage of forming chunks within a series to be maintained depended on temporal factors specific to the task at hand. A crucial temporal factor was the moment at which the chunk could be formed. We postulated that the earlier chunking arises within a series to be maintained, the greater the number of opportunities of refreshing one chunk instead of refreshing three independent elements, and hence the greater the benefit to be observed on recall performance. Behavioral results showed a superiority of acronyms in the first position compared with the third and fifth positions which did not differ from each other. Thus, the nice putative gradient of benefit increasing from later to earlier positions was not observed. A possible explanation could be made in reference to the well-known primacy and recency effects that boost performance in first and final positions, respectively (Murdock, 1962). These boosting effects observed in the present data and classically obtained in WM-CSTs (e.g., Oberauer & Lewandowsky, 2011; Portrat & Lemaire, 2015) could have

thus altered the postulated gradient by increasing even more the advantage of an acronym presented in the first positions but flattening, however, the difference between the middle and last positions.

The second temporal factor that could have influenced chunking benefit is the Pace of the concurrent activity. Its impact on the letter and acronym recall performance appeared to be more complex than anticipated. While chunking could be seen as a maintenance mechanism beneficial to WM performance, it is probably not competing with other attentional mechanisms in a simple manner. Indeed, the classic Pace effect favoring the slow pace was observed, as usual, in the condition with no acronym as well as in the conditions with acronyms whatever their position. According to the TBRS theory, this Pace effect marks a trade-off between attention-demanding processes that compete for the same limited amount of resource (e.g., Barrouillet et al., 2011). Thus, considering these behavioral outcomes, one could argue that chunking competes with other attentional processes, such as refreshing. However, considering now the TBRS*C parameters leads us to reconsider this simple trade-off conclusion between chunking and refreshing. Indeed, in the best model, chunking appeared to occur only during the encoding periods of the task. The model spends 1,800 ms after each new encoded letter to search for a chunk, by reactivating previously encoded letters and querying LTM for items chained together. On the contrary, the parameter controlling the process of searching for a chunk during the processing phase is set at zero. Hence, the Pace effect observed with acronyms is most probably not due to a competition between chunking and the concurrent processing activity. This interpretation is in turn confirmed by behavioral results showing that the Pace did not affect the recall of acronyms themselves, while on the contrary, the pace did have a classic effect on the isolated letters recalled in conditions with acronyms. The present outcomes give support to models using the idea that retrieving and using a chunk have a (time) cost, which competes with other processes that also have a time cost (Jones et al., 2007; Richman et al., 1995; Waters & Gobet, 2008). However, our results also specify that, depending on the time course of the

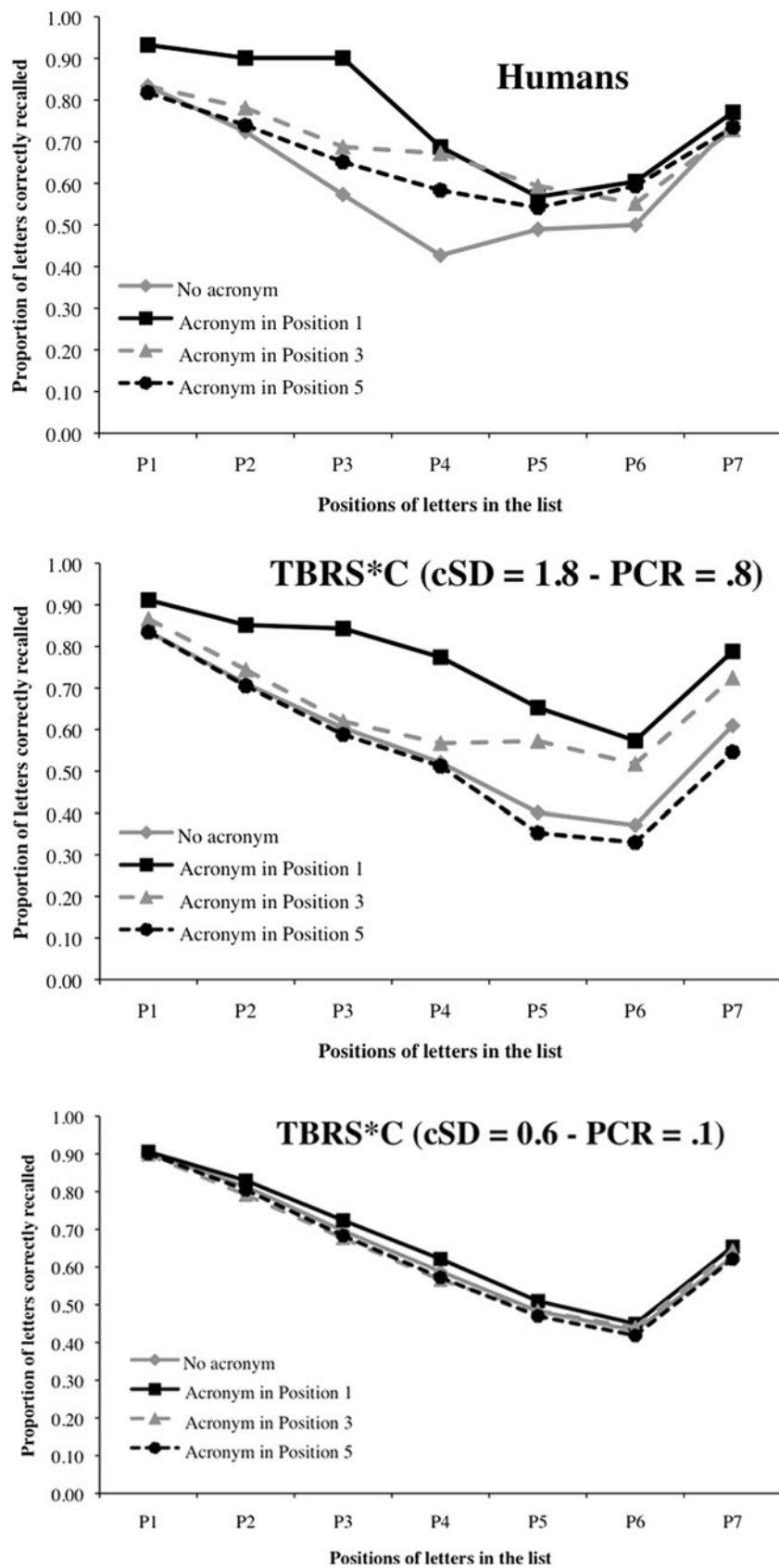


Fig. 8 The proportion of letters correctly recalled as a function of the serial position in the list (from P1 to P7) and the Acronym condition (no acronym, in Positions 1, 3, and 5) observed in humans (top panel),

produced by the model M1 (cSD = 1.800 ms, PCR = .8) (middle panel) and by the model M2 (cSD = 600 ms, PCR = .1) (bottom panel)

task at hand, chunking is not likely to be in competition with all the time-consuming processes. Chunking, a mechanism that links WM to LTM, seems to do so mainly at the precise moment when information has to be encoded. This outcome seems to be highly consistent with the widespread conception of WM as an activated part of LTM (e.g., Anderson, Reder, & Lebiere, 1996; Barrouillet et al., 2004; Cowan, 2005a; Just & Carpenter, 1992; Oberauer, 2002). Actually, these models have been called “LTM-based WM models” (Guida et al., 2013a, b). Encoding information in WM for short-term retention corresponds to reactivating LTM knowledge, this knowledge obviously includes groups of individual items.

Moreover, we also postulated that chunking is a maintenance mechanism that could be likened to refreshing. However, while certainly being a memory enhancer for the reasons described above, it should most probably not be considered as a maintenance mechanism per se. Two pieces of evidence run counter to such an account. First of all, the pace has no effect on the recall of acronyms. Hence, since the pace effect is a marker of a maintenance mechanism appearing during processing activity, chunking should not be considered as such. Moreover, according to the simulation, after chunking has taken place, the resulting chunk has to be maintained. TBRS*C refreshes chunks by attentional focusing able to re-activate four items (acronyms or single letters) at a time.

Finally, chunking appears to be a specific mechanism that is not similar to the other maintenance mechanisms at hand in WM. Chunking does not suffer from exactly the same constraints as the attentional refreshing described in TBRS. On the contrary, the idiosyncrasy of chunking is that it allows a solid bridge between WM temporary content and LTM permanent knowledge. The literature on expertise should be enlightening in this regard. It is important to acknowledge explicitly the difference between two different processes that are both often named chunking—chunk creation and chunk retrieval—but that do not intervene at the same moment during the learning process (Guida, Gobet, Tardieu, & Nicolas, 2012; Guida, Gobet, & Nicolas, 2013). Chunk creation takes place first and starts when novices execute a novel task several times, with no expertise. They start binding together separate elements. This is thought to occur in WM (Oakes, Ross-Sheehy, & Luck, 2006; Wheeler & Treisman, 2002) and more precisely in the focus of attention (Cowan, 1999, 2001, 2005a), as already seen. As practice progresses, the binding process is repeated numerous times, the same elements always being chunked together. It results in a chunk creation, which is stored in semantic LTM. Once chunks have been created and thus stored in LTM, the second phase can start: chunk retrieval. Once a chunk is simply retrieved, chunking can be considered as a process that would not impose substantial cognitive cost but would constitute an important memory enhancer. Our experimental paradigm could be considered as a microreproduction of the aforementioned chunking process

including two steps. On the one hand, even if acronyms are already known and stored in LTM, a chunk creation stage is mandatory because of the scattered presentation of memory items. This chunk creation stage (that corresponds to the *cSD* parameter of the model) seems to occur especially during the encoding of new items and is supported by attention. On the other hand, when the chunk is created, i.e., when multiple elements are reencoded as one element, its retrieval (for refreshing or recall) does not seem to be particularly demanding.

The present work is the first attempt to supplement the well-known TBRS model with a chunking mechanism able to provide predictions on a given experiment. TBRS*C has to be tested with more data, in order to confirm the default parameters that we suggested. Moreover, the listing view of the LTM component is not as sophisticated, nor as psychologically plausible, as it should be. Other models such as EPAM (Elementary Perceiver and Memorizer; Feigenbaum & Simon, 1984) and its variants including CHREST (De Groot & Gobet, 1996) should be taken as fruitful backgrounds to improve TBRS*C in future research. In these models, LTM is represented as a discrimination net that is an efficient way of inserting and retrieving items. Our next step is to improve our LTM component in order for it to account for the increase of knowledge/expertise. Currently, the LTM component is a limited pool of 24 chunks. The size of the pool could be varied in order to adapt to the knowledge of participants. Moreover, we intend to instantiate and manipulate factors that could influence the probability of accessing a chunk, such as the frequency of use of acronyms. For example, a score of frequency could be computed for each acronym through a survey proposed to a large sample of French speakers and then implemented into our model. Our final goal would be to simulate well-known features of expertise, such as false recognition (Arkes & Freedman, 1984; Baird, 2003; Castel, McCabe, Roediger, & Heitman, 2007), which, for the moment, the model cannot handle. This could be implemented by a search for chunks that could start before the presentation of the third letter. Doing so, the model would anticipate and, in some cases (that need to be determined) of wrong anticipation, an anticipated acronym could be recalled (because more active) instead of the letters presented.

To sum up, together, the data and simulations provided a better understanding of the maintenance mechanisms taking place in WM. We believe that the experimental protocol we used, especially asking participants whether they found acronyms, encouraged them to deliberately search and form chunks, which gives us more assurance that our approach really concerns goal-directed chunking. Of particular importance, the present results lend evidence to the functional characteristics of chunking that seems to be, especially in a WM-CST, an attentional time-based mechanism that certainly enhances WM performance but also competes with refreshing, especially during encoding. We believe that this last element is

important for the WM community, which is more familiar with chunking as a memory enhancer.

References

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.
- Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*, 221–256.
- Arkes, H. R., & Freedman, M. R. (1984). A demonstration of the costs and benefits of expertise in recognition memory. *Memory & Cognition*, *12*, 84–89.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Baird, R. R. (2003). Experts sometimes show more false recall than novices: A cost of knowing too much. *Learning and Individual Differences*, *13*, 349–355.
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*, 570–585.
- Barrouillet, P., 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., Lépine, R., & Camos, V. (2008). Is the influence of working memory capacity on high level cognition mediated by complexity or resource-dependent elementary processes? *Psychonomic Bulletin & Review*, *15*, 528–534.
- Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, *118*, 175–192.
- Bor, D., & Owen, A. M. (2007). A common prefrontal-parietal network for mnemonic and mathematical recoding strategies within working memory. *Cerebral Cortex*, *17*, 778–786.
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*, 457–469.
- Camos, V., Mora, G., & Barrouillet, P. (2013). Phonological similarity effect in complex span task. *Quarterly Journal of Experimental Psychology*, *66*, 1927–1950.
- Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, *39*, 231–244.
- Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin and Review*, *22*, 1029–1034.
- Castel, A. D., McCabe, D. P., Roediger, H. L., III, & Heitman, J. L. (2007). The dark side of expertise: Domain-specific memory errors. *Psychological Science*, *18*, 3–5.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.
- Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1235–1249.
- Cohen, J. D., MacWhinney, B., Flat, M., & Provost, J. (1993). Pyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*, 257–271.
- Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental store capacity. *Behavioral and Brain Sciences*, *24*, 87–185.
- Cowan, N. (2005a). *Working memory capacity*. Hove, England: Psychology Press.
- Cowan, N. (2005b). Working memory capacity limits in a theoretical context. In C. Izawa & N. Ohta (Eds.), *Human learning and memory: Advanced in theory and applications: The fourth Tsukuba International Conference on Memory* (pp. 155–175). Mahwah: Erlbaum.
- Cowan, N., & Chen, Z. (2009). How chunks form in long-term memory and affect short-term memory limits. In A. Thom & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 86–101). Hove: Psychology Press.
- Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial task: A logical sequel to Miller (1956). *Psychological Science*, *15*, 634–640.
- Cowan, N., Saults, J. S., & Elliott, E. M. (2002). The search for what is fundamental in the development of working memory. In R. Kail & H. Reese (Eds.), *Advances in child development and behavior*, *29*, 1–49.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.
- De Groot, A. D. (1978). *Thought and choice in chess*. Hague: Mouton Publishers (**Original work published 1946**).
- De Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess: Heuristics of the professional eye*. Assen: Van Gorcum.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (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 & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge: Cambridge University Press.
- Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science*, *8*, 305–336.
- Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: Where do they come from? *Journal of Experimental Child Psychology*, *110*, 469–479.
- Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four ... or is it two? *Memory*, *12*, 732–747.
- Gobet, F., Lane, P. C. R., Croker, S. C. H., Cheng, P., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Science*, *5*, 236–243.
- Guida, A., Gobet, F., Tardieu, H., & Nicolas, S. (2012). How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework. *Brain and Cognition*, *79*, 221–244.
- Guida, A., Gobet, F., & Nicolas, S. (2013a). Functional cerebral reorganization: A signature of expertise? Reexamining Guida, Gobet, Tardieu, and Nicolas' (2012) two-stage framework. *Frontiers in Human Neuroscience*, *7*, 590.
- Guida, A., Gras, D., Noel, Y., Le Bohec, O., Quaireau, C., & Nicolas, S. (2013b). The effect of long-term working memory through personalization applied to free recall: Uncovering the primacy effect enthusiasm. *Memory & Cognition*, *41*, 571–587.
- Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory & Cognition*, *35*, 1675–1684.

- Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, 4, 268–280.
- Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective & Behavioral Neuroscience*, 5, 339–361.
- Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*, 10, 853–873.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review*, 99, 122–149.
- Kane, M. J., Conway, A. R. A., Hambrick, Z. H., & Engle, R. W. (2006). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). Oxford, England: Oxford University Press.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 33, 1–64.
- Laird, J. E., Rosenbloom, P., & Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning*, 1, 11–46.
- Lewandowsky, S., & Farrell, S. (2011). *Computational modeling in cognition*. Thousand Oaks: SAGE.
- Mathy, F., & Feldman, J. (2012). What's magic about magic numbers: Chunking and data compression in short-term memory. *Cognition*, 122, 346–362.
- 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.
- Mora, G., & Camos, V. (2013). Two systems of maintenance in verbal working memory: Evidence from the word length effect. *PLOS ONE*, 7, 1–8.
- Murdock, B. B. (1962). Effect of a subsidiary task on short-term memory. *British Journal of Psychology*, 56, 413–419.
- Naveh-Benjamin, M., Cowan, N., Kilb, A., & Chen, Z. (2007). Age-related differences in immediate serial recall: Dissociating chunk formation and capacity. *Memory & Cognition*, 35(4), 724–737.
- Oakes, L. M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science*, 17, 781–787.
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 28, 411–421.
- Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: a computational implementation of the time-based resource-sharing theory. *Psychonomic Bulletin & Review*, 18, 10–45.
- Pashler, H. (1998). *The psychology of attention*. Cambridge: MIT Press.
- 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(6), 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.
- Portrat, S., & Lemaire, B. (2015). Is refreshing in working memory sequential? A computational modeling approach. *Cognitive Computation*, 7, 333–345.
- Raye, C. L., Johnson, M., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: a minimal executive function. *Cortex*, 43, 135–145.
- Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory with EPAM IV. *Psychological Review*, 102, 305–330.
- Stuart, G., & Hulme, C. (2009). Lexical and semantic influences on immediate serial recall: A role for redintegration. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 157–175). Hove: Psychology Press.
- Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science*, 21, 384–390.
- Waters, A. J., & Gobet, F. (2008). Mental imagery and chunks: Empirical and computational findings. *Memory & Cognition*, 36, 505–517.
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131, 48–64.