



HAL
open science

Role of Attention in the Associative Relatedness Effect in Verbal Working Memory: Behavioral and chronometric perspectives

Fiona-Laura Rosselet-Jordan, Marlène Abadie, Stéphanie Mariz-Elsig, Valérie
Camos

► **To cite this version:**

Fiona-Laura Rosselet-Jordan, Marlène Abadie, Stéphanie Mariz-Elsig, Valérie Camos. Role of Attention in the Associative Relatedness Effect in Verbal Working Memory: Behavioral and chronometric perspectives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2022. hal-03551372

HAL Id: hal-03551372

<https://hal.science/hal-03551372v1>

Submitted on 1 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Role of Attention in the Associative Relatedness Effect in Verbal Working Memory:

Behavioral and chronometric perspectives

Fiona-Laura Rosselet-Jordan, Marlène Abadie, Stéphanie Mariz-Elsig, & Valérie Camos

Department of Psychology, University of Fribourg, Switzerland

Running Head: Relatedness Effect in Working Memory

Word count: 13553 Words (main text), 2 Figures, 5 Tables

Author Note

This work was supported by the Swiss National Science Foundation (grant 100019_175960) to Valérie Camos. The study received research ethics approval from the institutional review board of the University of Fribourg. Lists of words and data are available on OSF at <https://osf.io/sqz52/>. We thank Kim Heriksen, Andrea Fasel, Johanne Weber and Jovana Savic for their help in testing children, as well as Cathy Frei and the Schools Centre du Mail de Neuchatel, Centre scolaire du Val-de-Ruz and La Condémine for permitting testing in schools. M. Abadie is now at Laboratory of Cognitive Psychology, Aix-Marseille University, France. Correspondence should be addressed to Fiona Rosselet-Jordan Département de Psychologie, Université de Fribourg, Rue P.A. de Faucigny, 2 1700 Fribourg, Suisse. E-mail: fionalaura.rosselet-jordan@unifr.ch Phone number: +4126 300 7667.

ABSTRACT

Long-term semantic memory (LTM) is known for affecting recall during working memory (WM) tasks. However, the way LTM intervenes in WM remains unknown. Moreover, the available findings are incongruent concerning how attention modulates the impact of LTM on WM. To examine this issue, the involvement of LTM representations in a complex span task was manipulated through variations of the associative relatedness of memory items, while the attentional demand of the concurrent task was varied. Children and young adults were also compared, because children are less efficient in using refreshing for maintenance than adults. Despite the impact of the three main factors on recall performance, which was better for related than unrelated words, with the low rather than the high demanding concurrent task and for adults than children, there was no interaction between associative relatedness and attentional demand, neither in children nor in adults. We replicated these results in a second experiment with a more attention-demanding concurrent task. Moreover, analyses of recall latency showed that adults were faster than children at recalling words and both age groups were faster for related (vs. unrelated) words, but there was no effect of the concurrent attentional demand on recall latency and no interaction. Finally, errors were mostly omissions and transpositions, both more prevalent under high concurrent attentional demand. The present findings suggest that the availability of attention does not modulate the effect of LTM on WM. We discuss how WM models can account for this finding and how LTM can act on WM functioning.

WORD COUNT: 250

KEY-WORDS: Working Memory; Long-term Memory; Associative relatedness; Attention; Refreshing; Maintenance

The contribution of semantic long-term memory (LTM) is well known for improving recall performance in working memory (WM) tasks (Hulme, Maughan, & Brown, 1991; Schweikert, 1993). However, the way in which LTM intervenes in a WM task remains a topic of debate (see Thorn & Page, 2009, for a review). The present study aimed at examining the role of attention as a potential moderator of the effect of semantic LTM on WM functioning. For this purpose, we manipulated both the availability of attention during WM maintenance and the associative relatedness of the memoranda. Associative relatedness, which depends on semantic LTM, is known for impacting the ease of retrieving information from LTM (Collins & Loftus, 1975; Nelson & Zhang, 2000), whereas attention is depicted in many models as having a central role in WM maintenance (Barrouillet et al., 2004; Barrouillet & Camos, 2015; Cowan, 2001; Engle, 2002). However, there are inconsistent findings in the current literature that support different theoretical conceptions of how attention moderates LTM impact on WM. Hence, the present study aimed at examining this question.

The role of attention in mediating semantic LTM impact on WM

Several studies have shown that factors known for improving retrieval from LTM, such as lexicality, familiarity, frequency and relatedness, also impact recall performance in WM tasks (Hulme et al., 1991; 2003; Loaiza, Rhodes & Anglin, 2015; Poirier, et al., 2011; Saint-Aubin, Ouellette, & Poirier, 2005; Saint-Aubin & Poirier, 1999). The past five years showed the emergence of studies dedicated to the role of attention in mediating the LTM impact on WM. Many, if not all, of them were triggered by the interest in the functioning of attentional refreshing. Attentional refreshing is conceived as one of the main mechanisms to maintain verbal information in WM (Camos, 2015, 2017). Among the WM maintenance mechanisms, the most studied one is verbal rehearsal that maintains verbal information through phonological traces by using of language processes. Beside this domain-specific maintenance mechanism, attentional refreshing is thought to use (so-called executive) attention to

reactivate any types of memory traces. This domain-general mechanism increases the activation of recently presented, encoded, or retrieved information to keep it in an accessible state from moment-to-moment, thereby preventing it to fade out from immediate awareness (Camos et al., 2018). Directing attention to memory representations boosts, prolongs, and strengthens their activation, leading to better recall performance for the refreshed relative to non-refreshed representations. Although there is some consensus on the definition of attentional refreshing, its functioning remains unclear (Camos et al., 2018), and to explore its links with LTM, several studies examined how LTM effects were modulated by variations on its availability.

Rose, Buschsbaum, and Craik (2014) reported that the level-of-processing effect, one of the best-known LTM effect (i.e., better retrieval of deeply processed rather than shallowly processed words) affects recall in a WM task if attention is distracted by a concurrent task, but not when items can be maintained in primary (short-term) memory (see also, Rose, Craik, & Buschsbaum, 2015). Other studies contrasted LTM effect in simple vs. complex span tasks, the difference between the two tasks being indicative of the involvement of refreshing. Indeed, authors like McCabe (2008) or Loaiza and McCabe (2012) have put forward that the difference between the simple and the complex span tasks relies on the fact that complex span tasks provide refreshing opportunities, but not simple span tasks. According to these authors, refreshing is thought to use pauses between processing episodes in complex span tasks to retrieve into the focus of attention information that was displaced in secondary (long-term) memory during the processing episodes. There are no refreshing opportunities in a simple span task since there is no concurrent task and information is immediately recalled. Within this framework, Loaiza, Duperreault, Rhodes, and McCabe (2015) tested the lexicality effect in a simple and an operation span task. While words were better recalled than non-words in both span tasks, the lexicality effect was reduced in the operation span task compared to the

simple span task. Similarly, Abadie and Camos (2018) reported a smaller word frequency effect in a complex than in a simple span task. The same interaction between word frequency and task has been previously reported by Engle et al. (1990), when comparing the frequency effect in a simple vs. an operation span task. Overall, these studies reported a reduction of the impact of LTM effects on WM performance in complex span task, a task in which there are more refreshing opportunities than in simple span task according to McCabe (2008).

As suggested by Rose et al. (2014), such an interaction can find an explanation within the primary-secondary memory framework (James, 1890; Unsworth & Engle, 2007a, 2007b), in which items are primarily maintained in primary memory. When the current task exceeds the capacity of the primary memory (e.g., when the items to be maintained are too numerous or when refreshing opportunities are reduced like in simple span tasks or through distraction of attention), memory items are displaced from primary memory into secondary memory, and they must be retrieved from LTM for further recall through a cue-dependent search. Consequently, effects known for impacting LTM should be particularly evident when recall relies on retrieval from LTM, i.e., when refreshing is prevented (Rose et al., 2014). This could explain the findings reported by Engle et al. (1990), Rose et al. (2014), Loaiza et al. (2015) and Abadie and Camos (2018). However, besides these studies that reported an interaction between a LTM effect and the availability of attention for maintenance, others failed to observe any interaction (Camos et al., 2019; Campoy, et al., 2015; Loaiza & Camos, 2018). Examining the concreteness effect (i.e., more imageable words are better recalled than less imageable words) in immediate recall tasks, Campoy, et al. (2015) varied the availability of attentional resources by asking participants to concurrently perform either a simple tapping or a highly demanding random tapping task. As expected, the increase in attentional demand reduced recall performance, but the concreteness effect was not moderated by the concurrent tasks. Similarly, Camos et al. (2019) showed that the frequency and lexicality effect were not

modulated by the attentional demand of the concurrent task in complex span tasks, although both LTM effects affected recall performance. These authors also showed that frequency and lexicality effects did not moderate refreshing speed, which supports some independence of refreshing functioning from LTM. Using different types of manipulation, Loaiza and Camos (2018) provided further evidence of the independence of refreshing from LTM. In complex span task, the presentation of semantic cues during the recall phase had a beneficial effect on word recall performance compared to phonological cues. However, this benefit of the semantic cues did not vary as a function of the attentional demand of the concurrent task or even the intention to learn. Indeed, when participants were submitted to surprise recall test, semantic cue benefit remained the same as when they were instructed to actively maintain words for recall.

As proposed by Camos et al. (2019), the absence of interaction between LTM effect and the availability of attention can find an account within Cowan's (1999) Embedded-Processes (EP) model. In this model, WM is defined as the currently activated part of LTM, WM being a subset of LTM. Among this subset, 3-4 chunks of information are maintained in a high state of accessibility by a limited capacity focus of attention (Cowan et al., 2005). In this framework, one can expect that factors known for facilitating the retrieval from LTM (what we call here LTM effects) would also ease the encoding of items in WM. In other words, items easily retrievable from LTM have a higher probability to be activated and to constitute the content of WM. However, performing a concurrent task would more or less distract attention depending on its nature. This would lead to forgetting memory items. To counteract forgetting, memory traces are refreshed or reactivated through a rapid scanning of the central component of WM (Vergauwe & Cowan, 2014; 2015), akin to the mechanism proposed by Sternberg (1966; see also, Sternberg, et al., 1978). The sequential scanning would be performed on the set of memory items that constitute the central capacity-limited component

of WM and would reactivate the memory traces of these items. As a consequence, LTM effects should impact recall performance, because they affect the probability to enter into WM, but LTM effects should not vary with the availability of attention, which should impair the efficiency of the scanning among the subset of items that were in WM.

To summarize, there is some discrepancy within the literature about the interaction between LTM effects and the availability of attentional refreshing, and two streams of findings can be depicted. While some studies reported that LTM effects were amplified when refreshing opportunities were reduced (Abadie & Camos, 2018; Engle et al., 1990; Loaiza et al., 2015; Rose et al., 2014), others failed to report any interaction between LTM effects and manipulation of attention (Camos et al., 2019; Campoy et al., 2015; Loaiza & Camos, 2018). These contradictory findings contribute to the fact that the functioning of attentional refreshing is still unclear (Camos et al., 2018), but also support different theoretical conceptions on how attention mediates LTM effects in WM based on either the primary-secondary framework or the embedded-processes model. Hence, the present study aimed to re-examine this question by testing how variations in attentional refreshing moderate the relatedness effect. Like the previously mentioned studies, we examined recall performance. However, to help clarifying this discrepancy, we seek through other measures of cognitive functioning indexes of the role of attention on LTM effect in WM.

Recall latency and errors as indexes of memory functioning

The most obvious index of memory functioning is recall performance, and as such, it is the most commonly studied. However, examining recall latency and recall errors can be adequate indicators of the way memory is functioning. For example, in free recall tasks, it has been shown that recall probability and recall latency are independent measures (Rohrer & Wixted, 1994), while recall latency and recall errors account for unique variance (Unsworth, 2009).

Recall latency in WM span tasks has been the subject of several papers (Cowan et al., 1992, 1994, 1998, 2003; Hulme, Newton, Cowan, Stuart, & Brown, 1999; Towse et al., 2008) after a seminal study by Cowan (1992). These papers resulted on a set of congruent findings. First, the duration between the recall signal and the utterance of the first recalled word is systematically longer than the recall latencies for the other words. Authors referred often to this duration as a preparatory pause during which participants recapitulate the word list before they start recalling the words. As a consequence, this latency should increase with the list length, something that is sometimes observed (Cowan et al., 1994, 1998), but not systematically (Cowan et al., 1992). Moreover, preparatory intervals can be more complex to interpret than other inter-word pauses as they can reflect not only the retrieval of the memory list, but also some articulatory planning needed to output responses (Jarrold, Hewes, & Baddeley, 2000). Second, recall latency for the serial positions after the first one also increases with list length (Cowan, 1992; Cowan et al., 1994, 1998, 2003), an increase that is stronger in younger children (Cowan et al., 1998). These inter-word pauses are considered as an index of retrieval processes that contribute to recall performance (Cowan, 1992; Towse & Cowan, 2004). According to Rose et al (2014), when recall is rapid, this suggests that words were “reported directly from the focus of attention” (p. 695) and recall did not involve retrieval processes from secondary memory. On the contrary, when recall is slow, this suggests that it “involved cue-dependent search and retrieval from LTM” (p. 695). In addition, it has been shown that differences in the size of the search set could also be responsible for increasing recall latencies with list length (Rohrer & Wixted, 1994). Finally, children are slower than adults (Cowan et al., 1994), which is considered as resulting from their slower scanning pace (Keating, et al., 1980). It can also be suggested that children have larger search sets than adults resulting in longer recall latencies and a greater impact of list length on latencies. Studies within the primary-secondary memory framework that examined individual differences in WM showed that individuals with low WM

capacity exhibited longer recall latencies due to larger search set size. (Unsworth, 2009).

Similar results can be expected in children as they have lower WM capacity compared to adults.

The impact of LTM effect on recall latency was rarely examined, and the two available studies used the lexicality effect (Cowan et al., 2003; Hulme et al., 1999). This effect impacts recall latency, with longer inter-word pauses to recall lists of non-words compared to lists of words, and whatever the serial position (i.e., first and other serial positions). These findings are explained by a LTM support to WM performance. Because familiar items (like words) have stronger lexical representations in LTM, they are more easily reconstructed than unfamiliar items (non-words) during the pauses between recalling successive items, leading to shorter inter-word pauses and better recall (Hulme, et al., 1999).

Concerning recall errors, they are not systematically reported in studies using WM span tasks. Nevertheless, in two experiments using lists of related words, Atkins and Reuter-Lorenz (2008) observed that semantic errors (i.e., recalling an unrepresented theme word or an associated word) exceeded other types of errors, in particular they occurred more often than phonological intrusions and other unrelated intrusions. These semantic errors were greater when participants had to complete a high-demanding concurrent task while maintaining memory words than when there was no concurrent task.

Except this study, when errors were examined in WM tasks, this was mostly in recognition test (e.g., Abadie & Camos, 2019; Flegal, Atkins, & Reuter-Lorenz, 2010; McBride, Coane, Xu, Feng, & Yu, 2019; Shulman, 1972). Because the nature of the test has a direct impact on the types of errors (for example, recall test elicited obviously more omissions than recognition test in which it is not possible), it is difficult to use previous data on recognition test to support the examination of errors in recall test. However, we can take advantage of the work done in a related task, the simple span task, for which errors were reported in several studies

using recall test (e.g., Murdock, 1976; Saint-Aubin, Ouellette, & Poirier, 2005; Tse, 2009; Tse, Li & Altarriba, 2011).

Well replicated findings are that transpositions (i.e., recalling a word in a different serial position than the one in which it was presented) and intrusions (i.e., recalling an item from another list), especially semantic intrusions (i.e., recalling semantically related extra-list item) are more frequent in related than in unrelated word lists (Murdock, 1976; Saint-Aubin, Ouellette, & Poirier, 2005; Tse, 2009). However, it should be noted that intrusions are rare, often less than 1% when lists were not only semantically related word lists (Tse et al., 2011). These findings were often considered as evidence for the item redintegration theory. This theory predicts higher number of order errors for related lists, because degraded memory trace of an item has higher probability to overlap with features of another item in related lists. Hence, recall errors, as recall latency, can give insights about WM functioning.

The present study

In the present study, we conducted two experiments aimed at examining the mediating role of attention on the associative relatedness effect in WM. Semantic relatedness affects recall performance in serial recall task, with semantically related word lists better recalled than semantically unrelated word lists (Poirier et al., 2011; Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999; Tse, 2009). Two effects can be distinguished among the semantic relatedness effects: categorical relatedness that links words belonging to the same semantic category (e.g. *apple-banana-pear*) and associative relatedness, which is characterized by a high probability of associations between two or more words despite the fact that they do not belong to the same semantic category (e.g., *rabbit-carrot-ear*; Tse, et al., 2011). This latter type of associations is mainly based on the recurrent use or appearance of these words together. Two words can be highly associated while being semantically dissimilar (e.g. *rabbit-carrot*), or weakly associated

and yet semantically similar (e.g. *carrot-cabbage*; Thompson-Schill, Kurtz, & Gabrieli, 1998). The present study focused on the latter type of associative relatedness.

In both experiments, we manipulated the impact of LTM by presenting lists of either associatively related or unrelated words. In Experiment 1, these word lists were introduced in a complex span task in which adults and 11-year-old children had to perform either a low (a color judgment task) or a high (a parity judgment task) attention-demanding concurrent task. These tasks have been previously used in WM tasks with adults and children to vary the concurrent attentional demand (e.g., Barrouillet et al., 2004; Oftinger & Camos, 2016, for studies in adults and children, respectively). Increasing the attentional demand of the concurrent task performed during the delay of retention reduces the availability of attention for maintenance purpose, and in particular for refreshing memory items. To ensure that the observed effects only rely on the use of attentional refreshing, we hindered the use of articulatory rehearsal by imposing a concurrent articulation during the retention interval. Indeed, articulatory suppression by a concurrent articulation prevents the use of this subvocal repetition by inner speech strategy (Baddeley, 1986), with no impact on attentional refreshing in adults (e.g., Camos, Lagner & Barrouillet, 2009) and children (e.g., Mora & Camos, 2015; Oftinger & Camos, 2016, 2017, 2018).

We compared two age groups, children and adults, because development is a particularly powerful, quasi-experimental, manipulation of ability level. Indeed, it has been shown that children have a less efficient use of refreshing (Barrouillet et al., 2009; Camos & Barrouillet, 2011; Gaillard, Barrouillet, Jarrold, & Camos, 2011; Portrat et al., 2009). This was especially supported by the fact that WM recall performance in children was less affected than in adults by the increase of concurrent attentional demand. Indeed, by increasing the attentional demand of the concurrent task in complex span task, it has been shown that the decrease in children's recall performance was smaller than in adults; the slope related attentional demand to recall

performance being steeper in adults and older children (14-year-old) than in younger children (8-, 10-, 12-years of age). Moreover, Barrouillet et al. (2009, Exp. 2) assessed the rate of refreshing (i.e., the number of items refreshed per second) of two groups of children, 8- and for 14-year-olds, in a complex span task. They showed that the rate of refreshing is twice as high for 14- than for 8-year-old children, which suggests a stronger or more efficient use of refreshing in older children. In a previous experiment, the same authors showed that 14-year-old children have the same pattern of performance as young adults, with a similar impact of variation in concurrent attentional demand on recall performance, suggesting that age 14 is an end point in the development of the refreshing mechanism. Hence, testing children would allow us to assess LTM effect in WM in conditions of poor use of attentional refreshing, and according to one of the previously described sets of findings, this should strongly promote LTM effect.

Experiment 2 aimed at replicating Experiment 1 findings with younger children, a stronger manipulation of concurrent attentional demand and a more precise measure of recall latency. In both experiments, in addition to recall performance, we also assessed recall latency and recall errors, because these measures can also reflect how retrieval in a WM task can be influenced by LTM knowledge.

As previously explained, we expected to replicate well-known effects of age, with poorer recall performance in children than adults, of associative relatedness with better recall performance for related word lists than unrelated word lists, and of concurrent attentional demand, its increase leading to reduced recall performance. This latter effect should be stronger in adults than in children, because the former are more efficient in using attention for maintenance purposes than the latter. However, the main purpose of the current study is related to other interactions.

The main point of interest of the present study concerned the effect of associative relatedness as a function of the availability of attentional refreshing, which can be impaired either experimentally through variations in attentional demand of the concurrent task or through the comparisons between age-groups as children spontaneously present a poorer use of refreshing. As previously described, the literature reported two streams of results: either LTM effect is increased under the impairment of refreshing or LTM effect is immune to it. Each of these patterns of findings found an account within either the primary-secondary memory framework or the embedded-processes model, respectively. Moreover, because children are less affected by variation in attentional demand, it can be extrapolated that, if the increase of LTM effect under high attentional demand, as expected by the primary-secondary framework, was observed, it should be greater in adults than in children. Finally, and although this was never reported in the literature, it remains possible that LTM effect would be enhanced under low concurrent attentional demand, as suggested by the time-based resource sharing model (Barrouillet & Camos, 2015).

Contrasting with the numerous studies on recall performance in WM tasks, there have been fewer investigations of recall times and recall errors. Previously observed findings should be replicated. Concerning recall latency, children should be slower than adults, the first recall latency should be longer than the others, latency should increase with list length and in unrelated word lists. Finally, the increase of latency with list length should be larger in children than in adults. Concerning recall errors, semantic errors should be more frequent in condition with high concurrent attentional demand and in related word lists, and the occurrence of order errors should increase in related compared to unrelated word lists.

Concerning the predictions issued from the two streams of findings and their supporting model, we can only speculate, as no explicit predictions have been made on recall latency and errors. Nevertheless, this helps in organizing the report of the findings. We can first expect

that the same interactions (or absence of interactions) predicted for recall performance should affect the other aspects of memory behavior, like latency and error. Second, the impact of concurrent attentional demand can also be envisioned. On the one hand, within the primary-secondary memory framework, Rose et al.'s (2014) suggested items retrieved from primary memory should then be retrieved faster than those in secondary memory. Therefore, recall latency should be longer under high attentional demand. On the other hand, the embedded-processes model would expect that concurrent attention-demand does not affect recall latencies, because items are recalled from the central component of WM in which items are in a high state of accessibility. Finally, intrusions, and in particular semantic intrusions, can occur more easily when items are retrieved from secondary than primary memory, which should be more frequent under high concurrent attentional demand, according to Rose et al.'s (2014) proposal. Our predictions and main results are summarized in Table 1.

EXPERIMENT 1

In this first experiment, we manipulated the impact of LTM by presenting lists of either associatively related or unrelated words. Following an increasing length procedure, these word lists were introduced in a complex span task in which adults and 11-year-old children had to perform either a low (a color judgment task) or a high (a parity judgment task) attention-demanding concurrent task. These tasks have been previously used in WM tasks with adults and children to vary the concurrent attentional demand (e.g., Barrouillet et al., 2004; Oftinger & Camos, 2016). To ensure that the observed effects only rely on the use of attentional refreshing, we hindered the use of articulatory rehearsal by imposing a concurrent articulation, i.e., continuously repeating three syllables aloud. As explained in the introduction, the role of attention in mediating the LTM impact on WM was examined in recall performance, as well as in recall latency and recall error.

METHOD

Participants

Participants included 33 undergraduate students of the University of Fribourg (28 females, $M_{\text{age}} = 20.9$ years, $SD = 1.83$) and 36 children from local schools (18 females, $M_{\text{age}} = 11.1$, $SD = .82$). All of them were French native speakers. They gave either a signed informed consent or a legal tutor agreement for children. Adults received course credits for participating. Children did not receive any reward for their participation, but they voluntarily participate to ensure their motivation. None of them took part in the pre-test of the material (see Supplementary Material).

Material

To our knowledge, no study in French ever created lists of several words to test the associative relatedness effect in WM tasks. Thus, we created our own lists of words using verbal association norms by De La Haye (2003) and by Duscherer, Khan, and Mounoud (2009). From these norms, 160 one to three syllable words were selected and divided into two sets of 80 words. In each set, we created four lists of associatively related words for each length (from two to six). The words were then mixed to create an equivalent number of lists of two to six unrelated words (see Supplementary Material for the pre-test of the material). Each participant saw each word only once in the experiment, either in a related or unrelated list. The related lists of one set and the unrelated lists of the other set were administered to each participant. For the concurrent tasks, series of blue and red digits ranging from one to eight were randomly generated with a rate of 50% of each color and 50% of each parity.

Procedure

The task was programmed on E-Prime 2.0 (Psychology Software Tools, 2012). Participants performed a complex span task during which they maintained lists of words while performing a concurrent task, either the color (low attentional demand) or parity (high

attentional demand) judgment task on series of digits (Figure 1). Each trial began with the presentation of an asterisk for 2000 ms indicating the start of the concurrent overt articulation. Throughout each trial, participants repeated aloud the syllables *ba-bi-boo*, the experimenter tracking their continuous utterance. This concurrent articulation hinders articulatory rehearsal (Camos et al., 2009; Cowan et al., 1998), and favors the use of attentional refreshing (Camos, Mora, & Oberauer, 2011). After the ready signal, the first to-be-remembered word appeared. Each word was presented during 1500 ms, followed by four digits presented for 1200 ms each with 300-ms inter-stimuli interval (ISI), during which a fixation cross kept participant fixating the screen. Word lists were presented in increasing length from two to six with four lists of related and four lists of unrelated words in each length. Within each length, the eight lists were randomly displayed, with the constraints that four lists were assigned to each concurrent task, with half of them being related word lists and the other half unrelated word lists. In a counterbalanced way, participants studied each length, starting either by the color or parity judgement task, and completed the secondary tasks by pressing on *l* (for red or odd digits) and *s* (for blue or even digits) keys on a Swiss keyboard. At the end of a trial, a question mark with a beep prompted the oral recall of words in their order of presentation. The experimenter took note of the recall on protocol sheets. Participants were also asked to press the spacebar on the keyboard each time they recalled a word to record recall latencies. Because complying with this latter instruction could be difficult for children, the experimenter was in charge of pressing the key each time a child recalled a word. The length of word lists as well as the concurrent task to be performed (color or parity judgment task) were indicated on screen at the beginning of each trial.

The experiment started with two training phases. The first one was dedicated to the color and parity judgment tasks performed each on 20 digits, after an example on three digits. The second training phase was analogous to the complex span task, using lists of forenames

instead of words. Participants performed four lists of length two, half with the color judgment task and half with the parity judgment task; the order of the two conditions being counterbalanced across participants. Finally, in a post-test at the end of the experiment, participants judged the associative relatedness of the word lists that they had not studied in the experiment (see Supplementary Material for the results of this post-test).

RESULTS

To control that participants complied with the instructions and paid enough attention to the concurrent tasks, data from participants with an accuracy rate lower than 65% were excluded. This led to exclude the data from two children. Moreover, data from participants with a span score differing from the average span of their age group by more than 3SDs were also excluded. This resulted in the exclusion of the data of one adult. The data from one additional child were excluded due to a computer error and another one refused to finish the experiment. Additional participants replaced these missing participants in order to have an equal sample size ($N=32$) in children and adults. The reported analyses were performed on this sample.

All analyses were performed with JASP version 0.11.1 (JASP Team, 2019). For each dependent variable, a Bayesian Analysis of Variance (ANOVA) was performed using the default settings. The BF_{10} of each model (e.g., main effects only, main effects + interaction effects) was obtained by comparing it to the null model. For each dependent variable, we first report the best model, i.e., the model with the largest BF_{10} . Then, we report the $BF_{inclusion}$ value for each factor in the model (i.e., a main effect or an interaction effect), which indicates the likelihood of the data under models that included a given factor compared to all models stripped of the factor. Table 2 provides the mean and SD for each dependent variable and Table 3 summarizes the $BF_{inclusion}$ of main effects and interactions. Before reporting the results

on recall performance, latency and error, we performed analyses on the concurrent task performance.

Performance in Concurrent Tasks

Percentage of correct responses (ACC) and reaction times (RTs) in the concurrent tasks were submitted to default Bayesian mixed ANOVA, with concurrent task (color vs parity judgement task) and relatedness (related vs. unrelated lists) as within-subject factors, and age group (adults vs. children) as a between-subject factor.

Concerning the percentage of correct responses, the best model included the main effects of concurrent task and age group and the interaction between the two factors ($BF_{10} = 3.57 \times 10^{36}$). As expected, the percentage of correct responses was higher in the color ($M = 91\%$, $SD = .06$) than in the parity judgement task ($M = 82\%$, $SD = .12$), and in adults ($M = 91\%$, $SD = .05$) relative to children ($M = 82\%$, $SD = .09$). Moreover, follow-up Bayesian paired t-tests showed that the impact of concurrent task was larger in children ($BF_{10} = 5.07 \times 10^6$) than in adults ($BF_{10} = 1551$).

Concerning RTs, the additive model including the main effects of the three factors was the preferred model ($BF_{10} = 1.27 \times 10^{41}$). As expected, participants were faster when performing the color ($M = 630$ ms, $SD = 76$) than the parity judgment task ($M = 709$ ms, $SD = 74$). They were also faster to perform the concurrent task when maintaining lists of related ($M = 661$ ms, $SD = 76$) than unrelated words ($M = 678$ ms, $SD = 74$). Unsurprisingly, children were slower than adults ($M = 704$ ms, $SD = 70$ vs. $M = 635$ ms, $SD = 80$, respectively).

Recall Performance

Recall performance was first scored as span, with span being equal to $1 + (1/4 \times N)$ with N the total number of trials in which all words were correctly recalled in their correct serial position (e.g., Barrouillet et al., 2009; Smith & Scholey, 1992). Because this measure may not

be sensitive enough to grasp small recall differences, recall performance was also scored as percentage of serial recall and of free recall, i.e., irrespective of serial position.

Bayesian mixed ANOVA with concurrent task and relatedness as within-subject factors and age group as between-subject factor indicated that the additive model including the main effects of the three factors was the best model to account for span ($BF_{10} = 7.26 \times 10^{33}$). As can be seen in Table 2 and as predicted, span was higher when participants performed the color ($M = 3.07$, $SD = .11$) rather than the parity judgment task as concurrent task ($M = 2.81$, $SD = .10$), when they maintained related rather than unrelated word lists ($M = 3.50$, $SD = .12$ and $M = 2.37$, $SD = .09$, respectively) and in adults ($M = 3.34$, $SD = .93$) than in children ($M = 2.54$, $SD = .92$). However, the analysis provided substantial evidence against the interactions of interest between relatedness and concurrent task, and between relatedness and age. There was also weak evidence against the interaction between the three factors.

A similar analysis was performed on the percentage of correct serial recall with one additional within-subject factor, the list length (2 to 6). As for span, the best model included the main effect of the four factors ($BF_{10} = 1.07 \times 10^{118}$). However, this model was only weakly preferred to the second-best model, which also included the four main effects and the interaction between relatedness and list length ($BF_{10} = 6.58 \times 10^{117}$). The effect of age group, relatedness, and concurrent task were similar as those in span (Tables 2 and 3). The percentage of serial recall decreased with list length ($M = 80\%$; $SD = 24.5$, $M = 71\%$; $SD = 25.2$, $M = 57\%$; $SD = 28.1$, $M = 49\%$; $SD = 26.7$ and $M = 41\%$; $SD = 25.8$ for list length 2 to 6, respectively). The post-hoc analyses showed that the list length effect was substantially stronger for the unrelated word lists (decreasing from 73 to 33% for lists 2 to 6) than for the related word lists (from 85 to 49%; $BF_{10} = 1.86 \times 10^{36}$).

The same analysis on the percentage of free recall led to a slightly different pattern of findings. The best model included the four main effects and interactions between concurrent

task and age group and between relatedness and age group, as well as interactions between list length and relatedness and list length and age group ($BF_{10} = 1.31 \times 10^{90}$). The four main effects were similar as those in span and serial recall analyses. Although the evidence for the concurrent task x age group interaction was weak (Table 3), we nevertheless conducted follow-up Bayesian paired t-tests in each age group. Evidence was gathered for an effect of concurrent task in children ($BF_{10} = 20.5$), while there was substantial evidence against this effect in adults ($BF_{10} = 0.19$). Concerning the relatedness by age group interaction, follow-up Bayesian t-tests indicated that the effect of relatedness was larger in children ($BF_{10} = 2.06 \times 10^{11}$) than in adults ($BF_{10} = 9.56 \times 10^8$). Moreover, as shown by another follow-up Bayesian paired t-tests, list length had no effect when recalling related words ($BF_{10} = .04$), but impacted recall of unrelated words ($BF_{10} = 1.72 \times 10^6$). Lastly, follow-up Bayesian ANOVA was performed in each age group showing that children were far more impacted by list length ($BF_{10} = 7.03 \times 10^9$) than adults ($BF_{10} = 1.91$). As in span and percentage of serial recall, there was substantial evidence against the interaction of interest between concurrent task and relatedness and also against the interaction including concurrent task, relatedness and age group.

Recall Latency

Recall latency was collected for each recalled word when recalled in correct serial position, and, for each participant, averaged per position in each condition for each length. Data diverging from more than 3 SDs from the mean (1.99% of the recall latencies) were replaced by the average of this condition for the same position and list length in the same age group. Moreover, because recall performance decreased across list lengths and to have a high percentage of RTs per presented words in adults as well as in children, we restricted our analyses to list lengths 2 to 4 (97%, 93%, and 80% in adults, and 90%, 84% and 69%, in children, respectively).

The Bayesian ANOVA with relatedness, concurrent task and serial positions (1 to 4 averaged across list lengths 2 to 4) as within-subject factors and age group as a between-subject factor showed that the best model included the main effects of relatedness, serial position and age group, as well as the serial position x age group interaction ($BF_{10} = 7.02 \times 10^{12}$). However, this model was only weakly preferred to the second-best model which did not include the relatedness effect ($BF_{10} = 3.71 \times 10^{12}$). Accordingly, there was weak evidence that related words were faster recalled than unrelated words ($M = 1933$ ms, $SD = 1009$ and $M = 2741$ ms, $SD = 1511$, respectively). Serial position had a decisive impact on recall latencies. Post-hoc comparisons indicated that participants took far more time to recall the first word ($M = 3535$ ms, $SD = 2454$) than the second ($M = 1368$ ms, $SD = 667$; $BF_{10} = 1.88 \times 10^6$), the third ($M = 1821$ ms, $SD = 307$; $BF_{10} = 3.26 \times 10^5$), and the fourth ($M = 2454$ ms, $SD = 2773$; $BF_{10} = 3.16$). Except these differences, there was weak evidence of differences between the other serial positions ($BF_{s10} < 1.69$). Although there was weak evidence that children ($M = 3926$ ms, $SD = 1145$) were slower than adults ($M = 1647$ ms, $SD = 403$) to recall words, the analysis showed decisive evidence for an interaction between serial position and age group. Bayesian independent t-tests provided decisive evidence that children ($M = 5689$ ms, $SD = 2741$) were slower than adults ($M = 2100$ ms, $SD = 417$) at recalling words at each position ($BF_{10} = 1.05 \times 10^5$, 5.75, and 14.75 for positions 1, 2 and 3, respectively), except for position 4 ($BF_{s10} = .95$). There was substantial evidence against the main effect of concurrent task, and the interactions between relatedness and concurrent task, and weak evidence against the interaction between relatedness, concurrent task and age group.

To complete the analyses of recall latencies, two additional Bayesian ANOVAs with relatedness, concurrent task and list length (2 to 4) as within-subject factors and age group as a between-subject factor were performed on recall latency for the first word and on the recall latency for the other words (i.e., averaged across positions 2 to 4). The analysis for the first

word indicated a preference for the model showing the main effects of relatedness and age group ($BF_{10} = 258$), whereas the same analysis for the subsequent words showed also the main effects of relatedness, age group, and list length, but highlighted an interaction between relatedness and age group ($BF_{10} = 3.64 \times 10^6$). For the subsequent words ($BF_{10} = 1042$), related words ($M = 1202$ ms, $SD = 748$) were faster recalled than unrelated words ($M = 1899$ ms, $SD = 1860$), and the same pattern was observed on the first word ($BF_{10} = 1.34$, $M = 2904$ ms, $SD = 1872$ and $M = 3492$ ms, $SD = 2783$, respectively). For both first and subsequent words, children were slower than adults ($M = 4366$ ms, $SD = 3976$ and $M = 2030$ ms, $SD = 679$, $BF_{10} = 1.62 \times 10^{13}$ vs. $M = 1783$ ms, $SD = 1853$ and $M = 1318$ ms, $SD = 755$, $BF_{10} = 9.78$, respectively). The preferred models showed that list length had no effect on recall latencies for the first words ($BF_{inclusion} = .08$), but did for subsequent ones ($BF_{inclusion} = 2275$). Indeed, the more the list length increased, the more the latencies for recalling the subsequent words increased ($M = 1313$ ms, $SD = 1283$ for length 2, $M = 1267$ ms, $SD = 944$ for length 3, and $M = 2071$ ms, $SD = 1685$ for length 4). Finally, for relatedness and age group interaction in subsequent word latencies, Bayesian independent t-tests showed that children ($M = 2291$ ms, $SD = 2780$) were slower than adults ($M = 1506$ ms, $SD = 941$) at recalling unrelated words ($BF_{10} = 7.99$), the age difference being smaller for the related words ($M = 1276$ ms, $SD = 925$ and $M = 1129$ ms, $SD = 570$, for children and adults, respectively; $BF_{10} = 8.69$). This interaction was not observed for the first words ($BF_{inclusion} = .44$). As in the previous analysis, there was substantial evidence against the main effects concurrent task ($BF_{inclusion} = .13$ and $.11$ for first and subsequent words, respectively), the interactions between relatedness and concurrent task ($BF_{inclusion} = .16$ and $.15$, respectively), and between relatedness, concurrent task and age group ($BF_{inclusion} = .19$ and 1.23 , respectively).

Recall Errors

Recall errors were classified into five types: transpositions, omissions, intrusions, phonological and semantic errors. Errors were defined as a transposition when a word was recalled but in a different position than the one it was presented (e.g., last presented word recalled as first), as an omission when a word was not recalled, and as an intrusion when the recalled word was studied in a previous list. When the recalled word differed from only one syllable or swapped two syllables from the original studied item and had no semantic link with the items in the list, it was classified as a phonological error. Recalled words having a semantic connection with the items presented in the list were classified as semantic errors. Based on these definitions, three independent judges classified errors. We used a very stringent criteria; when there was no full agreement between the judges on the error category, errors were treated as unclassifiable and were therefore excluded from the analysis. This represented only 2.15% of the errors, showing the high agreement rate between the judges.

Examination of Table 2 showed that omissions and transpositions were the two most frequent errors, while semantic, phonological and intrusion errors were quite rare. Although the low rates of the latter types of errors precluded statistical analysis, it is worth noting that semantic, phonological and intrusion errors exhibited specific patterns, showing that our three variables of interest (relatedness, concurrent task and age groups) impacted differently their production. Semantic errors were more frequent in related than unrelated lists, in color than parity task and in adults than in children. While both phonological and intrusions errors were less frequent in related than unrelated lists, the phonological errors were lower in color than parity task, and intrusions were conversely more frequent in color than in parity task. Finally, adults produced more intrusions than children, but their phonological error rates were quite similar.

For the omissions and transpositions, the preferred model in the Bayesian ANOVA with error types, concurrent task, and relatedness as within-subject factors, and age group as

between-subject factor included main effects of the four factors, and three two-way interactions (error types x relatedness, error types x age group, and relatedness x age group), as well as the interaction between error types, relatedness and age group ($BF_{10} = 2.24 \times 10^{39}$). Concerning the main effects, the percentage of omission and transposition errors was similar (see Table 3 for BFs). There was substantial to decisive evidence that participants made more errors in the parity ($M = 43\%$, $SD = 18$) than color judgment condition ($M = 38\%$, $SD = 17$, $BF_{inclusion} = 7.98$), when maintaining unrelated ($M = 48\%$, $SD = 18$) than related lists ($M = 33\%$, $SD = 17$), and that children ($M = 48\%$, $SD = 16$) produced more errors than adults ($M = 33\%$, $SD = 13$).

Concerning the two-way interactions, because the relation between their factors were qualified by a three-way interaction, we chose to decompose the higher order interaction between error types, relatedness and age group, for which we gathered strong evidence. Hence, we conducted analyses on the relatedness by age group interaction in each error type. There was weak evidence for the interaction in transpositions ($BF_{inclusion} = 2.66$) and substantial evidence in omissions ($BF_{inclusion} = 1065$). Children did more transpositions ($M = 23\%$, $SD = 12$, and $M = 18\%$, $SD = 9$, respectively, $BF_{10} = 20.25$) and less omissions ($M = 16\%$, $SD = 10$, and $M = 38\%$, $SD = 16$, respectively, $BF_{10} = 1.26 \times 10^{23}$) in related than unrelated words lists. The percentages of transpositions were not affected by relatedness in adults ($M = 20\%$, $SD = 11$ for related and $M = 21\%$, $SD = 9$ for unrelated word lists, $BF_{10} = .20$) while, as children, they did less omission in related ($M = 6\%$, $SD = 5$) than unrelated ($M = 19\%$, $SD = 10$) word lists ($BF_{10} = 1.56 \times 10^{15}$).

DISCUSSION

In summary, the pattern of findings was rather similar across the different recall scores (see Table 3). Our manipulations were successful in impacting recall performance in the expected way and independently of the scoring, with poorer recall in children than in adults,

in parity than color judgment task, and in unrelated than related word lists. However, whatever the score we examined, no evidence supported two of the main interactions of interest, that are the interactions between concurrent task and relatedness, and between the concurrent task, relatedness and age group. The interaction between age and relatedness was not supported, except when recall was scored without taking into account serial position (i.e., free recall score).

Findings in recall latency and error went in a similar direction, strengthening the results in recall performance, as none of the previously mentioned interactions received support. However, in response latency as well as in error, known findings were replicated. As expected, children were slower than adults, and recall latency increased with list length. As reported in previous analyses on response latency (e.g., Cowan et al., 1994, 2003), the first word of the lists took longer to be recalled than the following words. This could reflect the recapitulation of the entire word list before recalling each word in serial order. In addition, the shorter response latency for related word lists was weakly supported, indicating that as soon as memory words are maintained enough in WM to be recalled, the semantic links between these words does not speed up their retrieval. However, manipulation of the concurrent attentional demand did not affect response latency. These findings shed light on WM functioning, especially how refreshing is working and its relationships with LTM knowledge. These issues are discussed in the general discussion.

Concerning recall errors, transpositions and omissions were by far the most common errors, both in adults and children. Their rate of occurrence increased in a condition with a higher concurrent attentional demand. Unrelated word lists gave rise to more omissions, which shows that recall benefitted from the semantic links between the related words. This echoed what was observed in recall performance. Finally, transpositions were more frequent in related than unrelated word lists in children. This is congruent with previous findings

showing that associative relatedness between items increases the proportion of order errors (Saint-Aubin et al., 2005) In addition, Poirier et al. (2015) demonstrated that items for which the semantic activation was heightened by items within the same list are more likely to migrate toward earlier positions at recall. These findings suggest that associative relatedness can produce perturbations of representations of item order. Before discussing the theoretical implications of these findings, we reported a second experiment that aimed at strengthening our results.

EXPERIMENT 2

Experiment 2 has the same aim as Experiment 1, i.e., to examine the mediating role of attention on the relatedness effect in WM. However, three changes were implemented to strengthen our manipulations and measures. First, Experiment 2 involved younger children (aged 9) than those in Experiment 1. The lack of support for the expected interactions with age in Experiment 1 could result from smaller (than expected) age-related differences between 11-year-old children and young adults in the efficiency of refreshing. Hence, with younger children in Experiment 2, we aimed at increasing these age-related differences to assess whether or not we replicated the absence of interactions with age. Second, following the same idea, we implemented in Experiment 2 a stronger manipulation of the concurrent attentional demand by varying the pace of presentation of digits (either slow or fast) on which a parity judgement task was performed. Indeed, in Experiment 1, and contrary to what was previously observed, recall performance in adults was not more strongly affected than in children by the variation in concurrent attentional demand. We chose then another way to vary the concurrent attentional demand. It has been extensively shown that increasing the pace of presentation of distractors in complex span task has a detrimental effect on recall performance because it induces a stronger capture of attention needed for memory maintenance (e.g., Barrouillet et al., 2007; see Barrouillet & Camos, 2015, for a review).

Finally, we also improved our measures of recall latency by audio-recording the entire sequence of responses and measured the oral inputs based on the sound oscillographic display.

METHOD

Participants

Thirty-three students of the University of Fribourg (26 females, $M_{age} = 20.6$ years, $SD = 1.90$) and 32 children recruited in local schools (12 females, $M_{age} = 9.22$, $SD = .71$) took part in the experiment after providing either their informed consent or a legal tutor agreement for children. They were all native French speakers. As in Experiment 1, adults received course credits for participating and children participated voluntarily.

Material

Thirty-two new lists of four associatively related or unrelated words (16 lists each) were created following the same procedure as for the material in Experiment 1 (see Supplementary Material). The associative relatedness of the lists was pre-tested on 120 students at the University of Fribourg. Each list included one to three syllables neutrally valenced words, with a maximum of one verb per list.

Procedure

We used the same complex span task as in Experiment 1, except that concurrent task was always a parity judgement task for which the attentional demand was varied by changing the pace of presentation of the digits (Figure 2). Digits were all black and presented for 1700 ms in the slow pace condition and for 700 ms in the fast pace condition, both with a 300-ms ISI. Another change compared to Experiment 1 was that word lists were of a constant length of four. Finally, to improve the measure of recall latency, we used version 2.3.3 of Audacity®

recording and editing software¹ to record each participant's oral recall and extracted recall latency from oscillographic display on screen.

The training phase preceded the task was akin to the one in Experiment 1, except that in the second phase with the complex span task one list was presented for each pace condition.

RESULTS

We excluded data from participants with performance in the concurrent and recall tasks that differed from more than 3SDs of their age group's mean. This resulted in data exclusion for one adult due to poor performance in the concurrent task, and one adult and one child for poor recall performance. The following analyses were then performed on 31 adults and 32 children. Table 4 shows mean and SD for each dependent variable and Table 5 the $BF_{inclusion}$ value for each main effect and interaction.

Performance in Concurrent Task

Bayesian ANOVAs on percentage of correct responses and RTs with pace and associative relatedness as within-subject factors and age group as between-subject factor lead to the same preferred model, which included main effects of pace and age group and a pace by age group interaction ($BF_{10} = 1.57 \times 10^{52}$ and $BF_{10} = 2.55 \times 10^{60}$, respectively).

As expected, accuracy was higher and RTs longer in the slow ($M = 85\%$, $SD = 14$, and $M = 697$ ms, $SD = 135$, respectively) than in the fast pace condition ($M = 75\%$, $SD = 17.4$, and $M = 541$ ms, $SD = 52.2$, respectively). Adults ($M = 92\%$, $SD = 6.19$) outperformed children ($M = 68\%$, $SD = 15.4$) and they were also faster ($M = 580$ ms, $SD = 86.8$, vs. $M = 658$ ms; $SD = 151$, respectively). Follow-up Bayesian paired-t-tests performed in each age group revealed decisive effect of pace in both age groups, but with a larger pace effect in children ($BF_{10} = 1.07 \times 10^{10}$ and $BF_{10} = 1.85 \times 10^{11}$) than in adults ($BF_{10} = 1.10 \times 10^6$ and $BF_{10} = 353016$) both for accuracy and RTs.

¹ Audacity® software is copyright © 1999-2019 Audacity Team. The name Audacity® is a registered trademark of Dominic Mazzoni.

Recall Performance

Because we presented constant length trials, recall performance was scored as percentages of correct serial and free recall. The preferred model for percentages of serial recall included the three main effects of pace, relatedness and age group and the interactions pace by age group and relatedness by age group ($BF_{10} = 1.56 \times 10^{43}$). However, this model was only weakly preferred to the second-best model, which did not include the relatedness by age group interaction ($BF_{10} = 1.13 \times 10^{43}$).

As expected, recall performance was decisively better in the slow ($M = 54\%$, $SD = 11.9$) than in the fast pace condition ($M = 45\%$, $SD = 14.0$). There was also decisive evidence for the relatedness effect, with related words better recalled than unrelated words ($M = 55\%$, $SD = 12.7$ and $M = 43\%$, $SD = 13.3$, respectively). Moreover, there was decisive evidence that adults recalled more words ($M = 71\%$, $SD = 12.9$) than children ($M = 28\%$, $SD = 13.1$). Regarding the pace by age group interaction, follow-up paired t-tests conducted separately in each age group provided decisive evidence for the pace effect in adults ($BF_{10} = 45508$), but weak evidence against this effect in children ($BF_{10} = 0.47$). Although the evidence for the relatedness by age group interaction was weak, the follow-up t-tests indicated that the relatedness effect was stronger in children ($M = 36\%$, $SD = 14.1$, $M = 19\%$, $SD = 10.2$, for related and unrelated word lists respectively; $BF_{10} = 5.86 \times 10^6$) than in adults ($M = 77\%$, $SD = 9.82$, $M = 66\%$, $SD = 12.8$, respectively; $BF_{10} = 2794$). Finally, as in Experiment 1, the analysis provided weak evidence against the interaction between pace and relatedness and substantial evidence against the three-way interaction between pace, relatedness and age group.

Similar analyses were conducted on the percentage of free recall and confirmed the findings observed in percentage of serial recall with the same preferred model including the main effects of the three factors and the interactions pace by age group and relatedness by age

group ($BF_{10} = 1.46 \times 10^{65}$). The $BF_{inclusion}$ values for each main effects and interactions are given in Table 5.

Recall Latency

Participants' recall was audio-recorded, and recall latencies (in ms) were extracted from the oscillographic display between the recall signal (a beep) and the beginning of the first word for the first recall latency, and from the utterance of the previous word to the beginning of the following word for the other recall latencies. Recall latencies were collected for each word when recalled in correct serial position and, for each participant, averaged per position in each condition. Data from one adult and three children were excluded due to recording technical issues. The following analyses included then 30 adults and 29 children. Data diverging from more than 3 SDs from the mean of the condition in each position were replaced by the average of this position in this condition in the concerned age group (i.e., for 4.5% of the recall latencies).

The best model in the Bayesian ANOVA with associative relatedness, pace and serial positions (first to fourth) as within-subject factors and age group as between-subject factor highlighted the main effects of relatedness, serial position and age group, and the interactions between age group and serial position, as well as between relatedness and serial position ($BF_{10} = 3.40 \times 10^{17}$). However, this model was only weakly preferred to the second-best model, which did not include the relatedness by serial position interaction ($BF_{10} = 3.03 \times 10^{17}$). There was decisive evidence that related words ($M = 2590$ ms, $SD = 1697$) were faster recalled than unrelated words ($M = 3232$ ms, $SD = 2313$). Recall latencies were also decisively impacted by serial position. Indeed, post-hoc comparisons indicated that participants took far more time to recall the first word of the lists ($M = 3378$ ms, $SD = 2034$) than the second ($M = 2749$ ms, $SD = 1753$; $BF_{10} = 92.6$) and the fourth ($M = 2373$ ms, $SD = 1642$; $BF_{10} = 1.64 \times 10^6$), but not the third ($M = 3144$ ms, $SD = 2531$; $BF_{10} = 0.14$). There was

weak evidence that recalling the third word took more time than recalling the second ($BF_{10} = 1.61$) and decisive evidence that it took more time than recalling the fourth word ($BF_{10} = 767$). Finally, there was substantial evidence that the fourth word was faster recalled than the second one ($BF_{10} = 3.63$). As in Experiment 1, there was decisive evidence that children were slower ($M = 3390$ ms, $SD = 2020$) for recalling words compared to adults ($M = 2448$ ms, $SD = 1979$).

To examine the age group by serial position interaction, Bayesian independent t-tests compared age groups in each serial position. This provided substantial evidence that children and adults did not differ in recall latencies for recalling the first word ($M = 3323$ ms, $SD = 947$ and $M = 3432$ ms, $SD = 2067$, respectively; $BF_{10} = .272$). However, children were systematically slower than adults at recalling the following serial positions ($M = 3569$ ms, $SD = 1219$ and $M = 1956$ ms, $SD = 1062$; $BF_{10} = 42985$, for the second position, $M = 3711$ ms, $SD = 1482$ and $M = 2597$ ms, $SD = 2119$; $BF_{10} = 6.57$ for the third position and $M = 2957$ ms, $SD = 1073$ and $M = 1809$ ms, $SD = 1149$; $BF_{10} = 869$ for the fourth position, respectively). Although the evidence for the relatedness by serial position interaction was weak, we nevertheless conducted Bayesian analyses to test the relatedness effect in each serial position. Relatedness had no effect in the first and last serial positions ($BF_{10} = .37$, and $BF_{10} = .45$, respectively), but related words tended to be recalled faster than unrelated ones in the second and third serial position ($BF_{10} = 3.09$ and $BF_{10} = 12.9$, respectively). Akin to our results in Experiment 1, the analysis provided substantial evidence against the interaction between relatedness and pace, and between relatedness, pace, and age group.

Recall Errors

As in Experiment 1, we classified errors as transposition, omission, semantic, phonological and intrusion errors. Three independent judges did the classification. Only 3.18% of errors did not satisfy the classification criteria and were therefore excluded from the

analyses. As in Experiment 1, omissions and transpositions on which we performed statistical analysis were the most frequent errors (Table 4). Semantic, phonological and intrusion errors were rare with children producing more errors of each type than adults. As also observed in Experiment 1, semantic errors were more frequent in related than unrelated lists, while it was the reverse for phonological and intrusion errors. Finally, semantic and phonological errors were slightly more produced under the low than high pace task. It was the same for the intrusions in adults but the reverse for children.

A Bayesian ANOVA was conducted on the percentage of omissions and transpositions with error types, pace and associative relatedness as within-subject factors and age group as between-subject factor. The best model included the main effects of the four factors and the interactions, error types by relatedness, error types by age group, relatedness by age group and the three-way interaction between error types, relatedness and age group ($BF_{10} = 6.37 \times 10^{91}$). This model did not strongly differ from the second-best model, which included the same effects and interactions with the addition of the pace by age interaction ($BF_{10} = 3.55 \times 10^{91}$).

There was decisive evidence for a main effect of error types, omissions ($M = 31\%$, $SD = 19$) being more frequent errors than transpositions ($M = 12\%$, $SD = 10$). There was strong evidence that participants made more errors in the fast ($M = 24\%$, $SD = 11$) than in the slow pace condition ($M = 19\%$, $SD = 12$), decisive evidence that there were more errors in unrelated ($M = 25\%$, $SD = 12$) than related word lists ($M = 18\%$, $SD = 10$) and in children ($M = 33\%$, $SD = 5$) than adults ($M = 13\%$, $SD = 5$).

As in Experiment 1, because most of the two-way interactions were qualified by a three-way interaction, we chose to decompose the higher order interaction between error types, relatedness and age group by examining the relatedness by age group interaction in each error type. The pattern of findings was rather similar to what was observed in

Experiment 1. There was decisive evidence for the interaction between relatedness and age group in omissions and strong evidence for the interaction in transpositions ($BF_{inclusion} = 162.70$ and 32.90 , respectively). Both adults and children did less omissions in related ($M = 9\%$, $SD = 6$ and $M = 36\%$, $SD = 14$, respectively) than unrelated lists ($M = 26\%$, $SD = 11$, and $M = 62\%$, $SD = 12$, respectively), but the strength of evidence was larger in adults ($BF_{inclusion} = 6.58 \times 10^{19}$) than children ($BF_{inclusion} = 5.18 \times 10^{16}$). By contrast, both adults and children did more transpositions ($M = 12\%$, $SD = 6$, and $M = 22\%$, $SD = 17$, respectively) in related than unrelated word lists ($M = 6\%$, $SD = 5$, and $M = 10\%$, $SD = 12$, respectively), and the strength of evidence was larger in children ($BF_{inclusion} = 5.81 \times 10^6$) than adults ($BF_{inclusion} = 1.09 \times 10^5$).

DISCUSSION

To summarize, Experiment 2 replicated the pattern of findings obtained in Experiment 1 for recall performance, latency and error. Across the two types of recall performance scoring, we replicated that recall was better in the slow than fast pace condition, with related than unrelated word lists, and in adults rather than in children. As we expected, our manipulation of the concurrent attentional demand and the involvement of younger children did result on an age group by pace interaction, with the expected stronger pace effect in adults. We also observed a stronger relatedness effect in children than in adults. Despite these two interactions that gave better conditions for assessing our predictions than in Experiment 1, no evidence was gathered in favor of any of the two interactions of interest (relatedness by pace, and relatedness by pace by age group).

Similarly, for recall latency, results in Experiment 2 mimicked those in Experiment 1, with a longer time for the first than the following words, for children than adults, and for unrelated than related words. Nevertheless, the interactions of interest remained unsupported by the data. For recall errors, omissions and transpositions were still the most common error

types in both age groups. There were more omissions and transpositions in the fast than in the slow pace condition, and unsurprisingly, children made more errors than adults. Associative relatedness impacted the number of transpositions and omissions. Overall, the different manipulations had the expected effects, and many well-known effects were replicated, but nevertheless, the interactions of interest remained unsupported. This strengthened the findings of Experiment 1 in confirming the absence of such interactions in different aspects of recall. The general discussion exposed how these findings impact the current theoretical conceptions of WM maintenance, and of its relationships with LTM.

GENERAL DISCUSSION

The aim of this study was to investigate whether attention modulates the contribution of LTM knowledge in WM. The current state of affairs showed divergent findings that was interpreted among two distinct conceptions of WM, based either on the primary-secondary framework or the embedded-processes model. To vary the implication of LTM knowledge, we asked participants to maintain lists of either associatively related or unrelated words. We also manipulated the availability of attention to refresh memory traces by either implementing in complex span tasks concurrent tasks that differed in their attentional demand or by comparing two age groups, young adults and children, the latter being less efficient in refreshing. The different theoretical proposals were contrasted in their predictions and we analyzed the impact of the above-mentioned variables on recall performance, latency and error.

To summarize the main findings, our manipulations lead to the expected findings, replicating some well-known effects (see discussion of each experiment). When focusing on the predictions, the current findings are more in line with the previous studies that reported an absence of interactions between LTM effect and variation in attentional refreshing (Camos et al., 2019; Campoy, et al., 2015; Loaiza & Camos, 2018). In the present study, while the

increase in concurrent attentional demand and relatedness had the expected impact in recall performance, they did not interact with each other, in adults and children. This absence of interaction was also observed in recall latency, which was affected by the relatedness of the memory words, but not by the demand of the concurrent task. Finally, relatedness and concurrent attentional demand influenced the occurrence of the two most committed errors, omissions and transpositions. We discuss below the implications of these findings in light of the different WM models, and especially what the absence of interaction between associative relatedness and variation in the availability of attention implies for the role of attention in WM. Finally, we discuss how recall latency and errors can bring insights on WM functioning and in particular on its relationships with LTM. We discuss these two points in turn.

Attention and relatedness impact different WM components

Our results revealed a strong effect of associative relatedness on recall performance, both in children and adults. As LTM is well known for impacting WM retrieval, this relatedness effect was expected. Indeed, it is assumed that relatedness could intrinsically activate associated LTM parts and then increased the accessibility to these memory traces. LTM then contributes to the maintenance of WM traces by activating new information that could be retrieved from the items to be maintained. Encoding related items should then lead to superior retention performance. The manipulation of the concurrent attentional demand also showed a significant impact on performance, with a poorer recall in the high demanding condition. This result was also expected by any WM models that give to attention a major role in the maintenance of information in WM (e.g., Barrouillet & Camos, 2015; Cowan, 1995; Engle, 2002). As a consequence, any depletion of the attentional resources by a concurrent task would reduce their availability for maintenance purpose. The stronger demand of attention (by the parity judgment task and the fast pace condition in Experiments 1 and 2, respectively) led to a stronger reduction of attention for maintenance and to poorer recall

performance. Moreover, as expected, the effect of concurrent attentional demand was stronger in adults than in children in Experiment 2 in which children were younger and the manipulation of the concurrent attentional demand stronger. This replicates previous findings (e.g., Barrouillet et al., 2009) and indicates that adults were more efficient than children in using attention for maintenance purpose. The fact that the relatedness of memory items and the attentional demand of the concurrent task have both the expected effect on WM performance confirms the soundness of our manipulation. This was particularly important for the relatedness effect because, due to the absence of existing material in French, we had to create especially for this study the lists of memory items (see Supplementary Material).

The main interest of our study was to test the contrasted predictions concerning the interactions between LTM effect (associative relatedness) and variation in the availability of refreshing (through variations in concurrent attentional demand and contrasted age groups). The current findings are clear. Bayesian analysis strongly supported the absence of interactions. This departs from reported interactions between a LTM effect and refreshing availability (Abadie & Camos, 2018; Engle et al., 1990; Loaiza et al., 2015; Rose et al., 2014), and brings some support to an embedded-processes conception of the links between LTM and WM.

More specifically, our results are in line with the last suggestion on refreshing made within the embedded-processes model. Vergauwe and Cowan (2015) suggested that refreshing acts as the scanning of the central component of WM that reactivates the memory traces stored in this component. The idea that attentional refreshing is a reactivation has been favored by several authors (Barrouillet et al., 2004; Cowan, 1995; Johnson, 1992; McCabe, 2008; see Camos et al., 2018, for a review and discussion on the different conceptions about refreshing). Hence, several models that share the idea that memory traces are stored in a central buffer in which they can be reactivated via attention can account for the current

findings. Within Cowan's embedded-processes model, memory items stored in the focus of attention benefit from a heightened accessibility. Refreshing would act as a sequential scanning or search within this central component of WM, leading to their sustained maintenance. Such a description of the refreshing fits also with Oberauer's (2002) concentric model that distinguishes a single-item focus of attention within a capacity-limited region of direct access. In this model, refreshing would act as a rapid switching of the one-item focus of attention among the region of direct access. Similarly, in Barrouillet and Camos' time-based resource-sharing model, items encoded and maintained in the executive loop would be refreshed sequentially by attention. In the most recent description of Baddeley's multicomponent model, the authors accepted the idea that an attentional refreshing mechanism involves "the focusing of attention on the representation of material within the episodic buffer" (Baddeley, Hitch, & Allen, 2021). Overall, our finding strengthens this view, as the reactivation of WM traces would not depend on LTM, but on the availability of attention for maintenance. LTM effects would then affect WM recall performance by modulating either encoding or retrieval processes at recall.

Conversely, our results are for the most part at odds with the Rose et al.'s (2014) findings. According to Rose et al.'s (2014) proposal, LTM would be more strongly involved in WM recall in high attentional demand condition, because the to-be-remembered items would have a higher probability to be retrieved from secondary memory than from primary memory when attention is less available (but see Rose, 2020, for a modified version of this model). In line with this proposal, in Experiment 2 (and in Experiment 1 when recall is scored as free recall), the relatedness effect was stronger in children. This finding seems to contradict the overall pattern of results, which supports the absence of interaction between LTM and refreshing. However, if the relatedness effect was modulated by the availability of attention, it should have been stronger in adults than in children in the high attentional demanding

condition, because children were less affected by the variation of attention. This was not the case in any of our experiments. Hence, our results rather suggest that LTM is similarly involved in the retrieval of the memory items whatever the availability of attention. Nevertheless, our findings can be understandable within the primary-secondary memory model as originally depicted by Unsworth and Engle (2007). In this model, any concurrent task, whatever its attentional demand, would lead to the displacement of memory items in secondary memory, from which items are retrieved through a cue-dependent search. As a consequence, LTM knowledge are similarly involved as soon as a concurrent task is involved and whatever its attentional demand. In the case of the relatedness effect, maintaining semantically related items would provide more efficient cues (e.g., the common gist underlying the list of items) to search for the items in secondary memory, leading to the observed benefit in recalling the related items (see Poirier et al., 2011; Saint-Aubin & Poirier, 1999). This echoes Oberauer's (2009) study in which the nature of the distractors that were presented after a memory item was manipulated in a complex span task. The distractors were either semantically related (presenting *Silk*, *Wool*, *Velvet* and *Linen*, after *Cotton*) or unrelated (presenting *Queen*, *Duke*, *Prince*, and *Knight* after *Cotton*). Despite the fact that the semantically related distractors improved recall performance of the memory items, probably by giving more efficient cues to search of the items, this semantic relatedness effect was not affected by variations in the attentional demand of the concurrent task.

Insights from recall latency and error on LTM impact in WM

As exposed in the introduction, it is quite rare to examine recall latency and error to get insights about WM functioning. However, previous studies provided interesting findings (Cowan et al., 1992, 1994, 1998, 2003; Hulme, Newton, Cowan, Stuart, & Brown, 1999; Towse et al., 2008), which showed that this is an avenue of research to further investigate. Hence, and contrary to most previous studies dedicated to the question of the role of attention

in the LTM effect on WM, the present study included analysis of recall latency to inform us about the ease to retrieve items from memory. As suggested by Rose et al. (2014), items retrieved directly from primary memory would be faster to be recalled than items retrieved from secondary memory. Therefore, according to these authors' proposal for the moderating role of attention, a high concurrent attentional demand or a poor use of attention for maintenance (as in children) should lead to longer recall latency. Alternatively, within the embedded-processes model, recall latency of items retrieved from WM should not be sensitive to the concurrent attentional demand, which conversely determines whether the items are present or not within WM. The analysis of the duration of the recall latency showed three main findings.

First, the attentional demand of the concurrent task had no impact on recall latency. In line with the embedded-processes model, we assume that memory traces are retrieved from a central component of WM, and as such, any items stored in this component would be retrieved and recalled at the same speed. As put forward by the time-based resource sharing model, the availability of attention would affect maintenance and thus, whether the items are still stored in this central component (Barrouillet & Camos, 2015). As a consequence, manipulating the attentional demand of the concurrent task affects the number of items stored (Barrouillet & Camos, 2014), and for these items, retrieval would be similar. Second, associative relatedness had a great impact on recall latency, related word lists being faster recalled. Accordingly, previous studies showed that LTM knowledge actually impacts recall latency, the retrieval of non-words producing longer inter-words pauses during recall than words (Hulme et al., 1999). This can be conceived like priming effects in LTM: retrieving and recalling the first item of a list could provide a kind of boost to utter the remaining items of the list when items are linked to each other. In this case, recall latency for the first item should not differ with relatedness conditions, and only the following items would show a beneficial

effect of relatedness. However, it is not what we observed in Experiment 1, in which relatedness weakly impacted the first four positions. In Experiment 2, we observed the predicted pattern with a relatedness effect occurring only in second and third positions, but not on the first position. Although this seems to be a divergence between our two experiments, it should be noted that there was evidence against the relatedness by serial position interaction in Experiment 1 and support for this interaction in Experiment 2 was weak. Hence, there might be another reason why relatedness impacts recall latency. One can envision that related items have stronger memory traces because they benefit from a semantic network in LTM (Ralph et al., 2017). Finally, congruently with Cowan et al. (2003), children took more time before recalling the first word in a list than adults in Experiment 1. We interpreted this effect as evidence of longer recapitulation time in children. Participants tend to recapitulate the word list before uttering the words. Children need then more time to build the list due to the time to retrieve each word, the larger size of the search set among which words have to be retrieved, or the difficulty in organizing their utterance. Indeed, Cowan and collaborators (1994, 1998) proposed that recall times provides an index of memory search. To prepare a response, a participant has to scan through all the stored items like in a Sternberg (1966) task in which individuals determine whether an item was previously presented, children being slower to do so. It can also be envisioned that children have larger set of candidates among which they have to search for the words to recall (Rohrer & Wixted, 1994; Unsworth, 2009, for studies in adults). Alternatively, Jarrold et al. (2000) proposed that these times reflect speech planning processes needed to utter the word list. However, this effect was not replicated in Experiment 2 and more importantly, recall latencies in both experiments were longer in children than in adults for subsequent positions. Although further studies are needed to examine what trigger these age-related differences, a simple explanation in terms of processing speed differences could account for this finding and was often mentioned as a

source of WM development (e.g., Camos & Barrouillet, 2018; Gaillard et al., 2011).

In addition to recall performance and latency, we analyzed recall error. Despite the fact that participants mostly omitted items when recalling lists of words, the impact of the relatedness on errors enlightens the mechanisms that sustain maintenance of verbal information at short term, and especially the role played by LTM. Indeed, participants produced less omissions when they had to recall related than unrelated words. This suggests that relatedness has a protective effect against forgetting. Related words may have stronger memory traces that better resist from representation-based interference and temporal decay. Relatedness also descriptively reduced the number of intrusions, that is the production of words from a previously presented list. When participants retrieved a word presented in previous trials, this word could be more easily rejected as a candidate for recall for lists of related words because these words share a common gist that differ from the gist of the other word lists. For example, it would be easier to reject the word "car" when the just-studied list is *Rabbit, Ear, Carrot* in the related condition than after the list *Rabbit, Line, Sport* in the unrelated condition). On the contrary, relatedness did not moderate the production of phonological errors, which can be expected as it would depend on the phonological similarity between the memory words. However, we were expected that semantic errors would be enhanced by the presentation of related words, because the activation of a common gist would automatically activate other related words, as observed in the emergence of false memory in LTM (Reyna & Brainerd, 1995). This was not the case, but the rate of intrusions, phonological and semantic errors was so low that further studies should be dedicated to examine how false memory could emerged in short-term recall tasks, and what is triggering the production of errors.

Finally, it should be noted that the increase in concurrent attentional demand resulted on more omissions. This was particularly expected by the time-based resource sharing model

because the number of items maintained in the episodic buffer is a direct function of the attentional demand of the concurrent task, that is the amount of time during which attention is available for refreshing activity (Barrouillet & Camos, 2015). Increasing the attentional demand of the concurrent task directly reduces the availability of attention for maintenance, resulting in a smaller number of maintained items.

CONCLUSION

The relationships between LTM and WM has been and still is source of intense theoretical debates in the literature. Although it is well known that knowledge stored in semantic LTM improves recall performance from WM, the exact pathway of such an effect remains under investigation. Within this line of research, the present study proposed to examine the moderator role of attention on the effects that semantic LTM had on WM functioning. In the past five years, an interest on this specific question emerged to highlight the functioning of one particular WM maintenance, attentional refreshing. This leads to a rather divergent pattern of findings on how attention may or not moderate LTM effects in WM.

In the present study, we have been able to replicate the beneficial effect of associative relatedness between memory items and the availability of attention for maintenance on WM performance. However, the vast majority of our pieces of evidence tends to support an independent impact of these two factors on WM functioning. This finding is in line with previous studies (Camos et al., 2019; Campoy, et al., 2015; Loaiza & Camos, 2018), and brings some support to a class of WM models in which attention is used to sequentially reactivate memory traces stored in a central component of WM (Baddeley et al., 2021; Barrouillet et al., 2004; Cowan, 1995; Oberauer, 2002). Moreover, from an age-related perspective, the present study showed that children exhibited rather similar pattern of findings than adults, besides expected poorer performance. This could evidence that children's use of

LTM knowledge to support WM did not qualitatively differ from adults, contrary to what was reported in other types of tasks. Finally, we believe that examining recall latency and error helped uncovering the role played by LTM in WM, and could be more frequently examined when studying WM.

References

- Abadie, M., & Camos, V. (2018). Attentional refreshing moderates the word frequency effect in immediate and delayed recall tasks. *Annals of the New York Academy of Sciences*, *1424*, 127-136.
- Abadie, M., & Camos, V. (2019). False memory at short and long term. *Journal of Experimental Psychology: General*, *148*(8), 1312-1334.
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK : Clarendon Press.
- Baddeley, A. D., Hitch, G. & Allen, R. (2021). A multicomponent model of working memory. In R. H. Logie, V. Camos, and N. Cowan (Eds). *Working Memory: State of the Science* (pp. 10-43). Oxford, UK : Oxford University Press.
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*(1), 83-100.
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*(3), 570-585.
- Barrouillet, P., & Camos, V. (2014). On the proper reading of the TBRS model: reply to Oberauer and Lewandowsky (2014). *Frontiers in Psychology*, *5*(1331), 1-3.
- Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove, UK : Psychology Press.
- Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development : A Time-Based Resource-Sharing model account. *Developmental Psychology*, *45*(2), 477-490.

- Camos, V. & Barrouillet, P. (2011). Developmental Change In Working Memory Strategies: From passive maintenance to active refreshing. *Developmental Psychology*, 47(3), 898-904.
- Camos, V., & Barrouillet, P. (2018). *Working memory in development*. Hove, UK: Routledge.
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, 61(3), 457-469.
- Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, 39(2), 231-244.
- Camos, V., Mora, G., Oftinger, A.-L., Mariz-Elsig, S., Schneider, P., & Vergauwe, E. (2019). Does semantic long-term memory impact refreshing in verbal working memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 45(9), 1664-1682.
- Campoy, G., Castellà, J., Provencio, V., Hitch, G., & Baddeley, A. (2015). Automatic semantic encoding in verbal short-term memory: Evidence from the concreteness effect. *Quarterly Journal of Experimental Psychology*, 68(4), 759-778.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.
- Cowan, N. (1999). An embedded-processes model of working memory. In Shah, P. and Miyake, A (eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge, UK: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87-185.

- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, 26(2), 197-223.
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conwan, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1), 42-100.
- Cowan, N., Keller, T., Hulme, C., Roodenrys, S., McDougall, S., & 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(2), 234-250.
- Cowan, N., Towse, J. N., Hamilton, Z., Saults, J. S., Elliott, E. M., Lacey, J. F., Moreno, M.V., & Hitch, G. J. (2003). Children's working-memory processes: A response-timing analysis. *Journal of Experimental Psychology: General*, 132(1), 113-132.
- Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General*, 127(2), 141-160.
- De La Haye, F. (2003). Normes d'associations verbales chez des enfants de 9, 10 et 11 ans et des adultes. *L'Année Psychologique*, 103(1), 109-130.
- Duscherer, K., Khan, A., & Mounoud, P. (2009). Recueil d'associations verbales chez des enfants de 5 à 11 ans pour 76 verbes d'action en langue française. *Swiss Journal of Psychology*, 68(2), 113-117.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(19), 19-23.
- 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(3), 469-479.

- Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, 39(1), 3-27.
- Gathercole, S. E., Pickering, S., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40(2), 177-190.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short term memory span. *Journal of Memory and Language*, 30(6), 685-701.
- Hulme, C., Newton, P., Cowan, N., Stuart, G., & Brown, G. D. A. (1999). Think before you speak: Pauses, memory search, and trace reintegration processes in verbal memory span. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25(2), 447-463.
- Hulme, C., Roodenrys, S., Schweikert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a reintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23(5), 1217-1232.
- Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, 49(4), 500-518.
- James, W. (1890). *The principles of psychology*. New York, NY: Holt and company.
- Jarrold, C., Hewes, A. K., & Baddeley, A. D. (2000). Two separate speech measures constrain verbal short-term memory in children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1626-1637.
- Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, 4(3), 268-280.

- Kowialiewski, B., & Majerus, S. (2018). The non-strategic nature of linguistic long-term memory effects in verbal short-term memory. *Journal of Memory and Language, 101*, 64-83.
- Loaiza, V. M., & McCabe, D. P. (2012). Temporal-contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition, 40*(2), 191-203.
- Loaiza, V.M., Rhodes, M.G., & Anglin, J. (2015). The influence of age-related differences in prior knowledge and attentional refreshing opportunities on episodic memory. *Journals of Gerontology, Lists B: Psychological Sciences and Social Sciences, 70* (5), 729-736.
- McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language, 58*(2), 480-494.
- Mora, G., & Camos, V. (2015). Dissociating Rehearsal and Refreshing in the Maintenance of Verbal Information in 8-Year-Old Children. *Frontiers in Psychology, 6*(11), 1-10.
- Nelson, D. L., & Zhang, N. (2000). The ties that blind what is known to the recall of what is new. *Psychonomic Bulletin & Review, 7*(4), 604-617.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUETM // A lexical database for contemporary french: LEXIQUETM. *L'Année Psychologique, 101*(3), 447-462.
- 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. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition, 37*(7), 346-357.

- Oftinger, A.-L., & Camos, V. (2016). Maintenance mechanisms in children's verbal working memory. *Journal of Educational and Developmental Psychology, 6*(1), 16-28.
- Oftinger, A.-L., & Camos, V. (2017). Phonological Similarity Effect in Children's Working Memory: Do Maintenance Mechanisms Matter? *Journal of Child Psychology, 1*(1), 5-11.
- Oftinger, A.-L., & Camos, V. (2018). Developmental improvement in strategies to maintain verbal information in working memory, *International Journal of Behavioral Development, 42*(2), 182-191.
- Poirier, M., Dhir, P., Saint-Aubin, J., Tehan, G., & Hampton, J. (2011). The influence of semantic memory on verbal short-term memory, In B. Kokinov, A., Karmiloff-Smith, & N.J. Nersessian (Eds.). *European Perspectives on Cognitive Science*. Sofia, Bulgaria: New Bulgarian University Press
- Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology, 48*(2), 384-404.
- Portrat, S., Camos, V., & Barrouillet, P. (2009) Working memory in children : A time-related functioning similar to adults. *Journal of Experimental Child Psychology, 102*(3), 368-374.
- Psychology Software Tools, I. E.-P. (2012). [E-Prime 2.0]. Retrieved from <http://www.pstnet.com>.
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*(1), 42-55.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1-75.

- Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition*, 22(5), 511-524.
- Rose, N. S. (2020). The dynamic processing model of working memory. *Current Directions in Psychological Science*, 29(4), 378-387.
- Rose, N. S., Buschsbaum, B. R., & Craik, F. I. M. (2014). Short-term retention of a single word relies on retrieval from long-term memory when both rehearsal and refreshing are disrupted. *Memory & Cognition*, 42(5), 689-700.
- Rose, N. S., Craik, F. I. M., & Buschsbaum, B. R. (2015). Levels of processing in working memory: differential involvement of frontotemporal networks. *Journal of Cognitive Neuroscience*, 27(3), 522-532.
- Rosselet-Jordan, F. L., Abadie, M., Mariz-Elsig, S., & Camos, V. (2021, March 24). Role of Attention in the Associative Relatedness Effect in Verbal Working Memory: Behavioral and chronometric perspectives. Retrieved from osf.io/sqz52
- Saint-Aubin, J., Ouellette, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: Is there an effect on all trials? *Psychonomic Bulletin & Review*, 12(1), 171-177.
- Saint-Aubin, J., & Poirier, M. (1999). The influence of long-term memory factors on immediate serial recall: An item and order analysis. *International Journal of Psychology*, 34(5-6), 347-352.
- Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schweikert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, 21(2), 168-175.
- Smyth, M.M., & Scholey, K.A. (1992). Determining spatial span: The role of movement time and articulation rate. *Quarterly Journal of Experimental Psychology*, 45A(3), 479–501.

- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652-654.
- Tehan, G. (2010). Associative relatedness enhances recall and produces false memories in immediate serial recall. *Canadian Journal of Experimental Psychology*, *64*(4), 266-272.
- Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, *38*(4), 440-458.
- Thorn, A., & Page, M. (2009). *Interactions between short-term and long-term memory in the verbal domain*. Hove, UK: Psychological Press.
- Towse, J. N., & Cowan, N. (2004). Working memory and its relevance for cognitive development. In W. Schneider, R. Schumann-Hengsteler, and B. Sodian (Eds.), *Young children's cognitive development : interrelationships among executive functioning, working memory, verbal ability and theory of mind* (p. 9-37). Mahwah, NJ : Lawrence Erlbaum Associates.
- Towse, J. N., Cowan, N., Hitch, G. J., & Horton, N. J. (2008). The recall of information from working memory: insights from behavioural and chronometric perspectives. *Experimental Psychology*, *55*(6), 371-383.
- Tse, C.-S. (2009). The role of associative strength in the semantic relatedness effect on immediate serial recall. *Memory*, *17*(8), 874-891.
- Tse, C.-S., Li, Y., & Altarriba, J. (2011). The effect of semantic relatedness on immediate serial recall and serial recognition. *The Quarterly Journal of Experimental Psychology*, *64*(12), 2425-2437.
- Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall.

Journal of Experimental Psychology: Learning, Memory and Cognition, 33(6), 1020-1034.

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition*, 37(6), 837-849.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104-132.

Vergauwe, E., Barrouillet, P., & Camos, V. (2014). The impact of storage on processing : How is information maintained in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 40(4), 1072-1095.

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychonomic Bulletin & Review*, 22(4), 1001-1006.

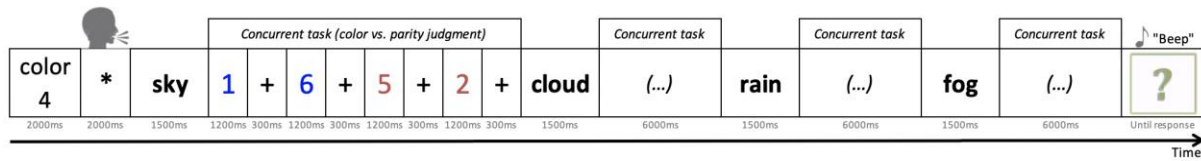


Figure 1. Example of a trial with a related list of four words in Experiment 1 in which participants had to remember the word list, while performing either a color-judgment (low demanding) or a parity judgment (high demanding) concurrent task. Each trial began by an indication of the concurrent task (color or parity) to be performed on digits and the number of memory words to be maintained. Starting at the ready signal, participants repeatedly uttered *ba-bi-boo* until the appearance of a question mark and a beep prompted aloud recall of the words in the order of presentation.

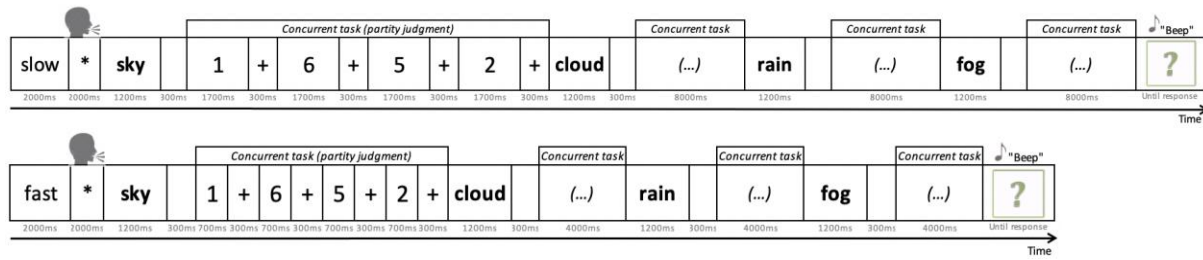


Figure 2. Example of two trials with a list of four related words in Experiment 2. Participants had to remember the word list, while performing a parity judgement task on digits that were presented either at a slow or fast pace. At the beginning of the trial, an indication mentioned the pace of the concurrent task. Starting at the ready signal, participants repeatedly uttered *ba-bi-boo* throughout the trial. Prompted by a question mark and a beep, participants orally recalled the words in the order of presentation at the end of each trial.

Table 1: Summary of predictions and main results

DVs	Replication of known effects		Based on primary-secondary memory framework		Based on embedded-processes model	
Recall Performance	Age: Poorer performance in children than in adults	✓	A stronger relatedness effect in children than in adults	(✓)	No interaction between relatedness and age	(✓)
	Concurrent attentional demand: Poorer performance in high than low attentional demand	✓	A stronger relatedness effect in high concurrent attentional demand than in low demand	x	No interaction between relatedness and concurrent attentional demand	✓
	Relatedness: Poorer performance in unrelated than related word lists	✓	A stronger increase of the relatedness effect in high (vs low) concurrent attentional demand in adults than in children	x	No 3-way interaction	✓
	Interaction Concurrent attentional demand by Age: a stronger effect in adults than in children	(✓)				
Recall Latency	Age: Children are slower than adults	✓	Age: Longer latency in children due to their poor use of refreshing or to larger search set size	✓	Age: Longer latency in children due to their slower scanning pace	✓
	Serial position: Longer latency for first serial position than the others	✓	Concurrent attentional demand: Longer latency in high than in low attentional demand	x	No effect of concurrent attentional demand	✓
	List length: Increased latencies with list length	(✓)				
	Relatedness: Longer latency in unrelated than related word lists	✓				
	Interaction Age by List length: A stronger increase with length in children than in adults	x	Interaction Age by List length: A stronger increase with length in children than in adults due to larger search set size	x		
Recall Errors	Relatedness: More semantic and transpositions in related than unrelated word lists	(✓) ✓				
	Concurrent attentional demand: More semantic errors in high than low attentional demand for related word lists	x				
	<i>Concurrent attentional demand: More omissions and transpositions in high than low attentional demand</i>	✓ ✓	Concurrent attentional demand: More intrusions in high than low attentional demand	*		
	<i>Relatedness: More omissions in unrelated than related word lists</i>	✓				

Note: Effects in Italics are not predictions based on the literature but observed in the current study. ✓ indicates that the effect was observed in both experiments, (✓) indicates that the effect was observed in one of the two experiments or descriptively present in recall errors, x indicates that the effect was not observed, and * indicates an inconclusive effect.

Table 2. Descriptive statistics of all dependent variables (DVs) in Experiment 1 as a function of age group (adults vs. children), the concurrent task (color vs. parity judgment task) and associative relatedness of word lists (related vs. unrelated words).

DVs		Adults				Children			
		Color		Parity		Color		Parity	
		Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
Recall	Span	4.0 (0.9)	2.9 (0.8)	3.8 (1.1)	2.69 (0.9)	3.2 (1.1)	2.2 (0.9)	3.1 (1.1)	1.7 (0.7)
	Serial recall	73.4 (15.3)	56.7 (17.2)	68.7 (17.2)	51.10 (20.5)	60.4 (19.4)	41.1 (20.5)	53.8 (21.5)	36.1 (17.4)
	Free recall	90.2 (7.4)	75.8 (12.6)	90.6 (8.5)	75.10 (13.5)	82.1 (10.0)	61.6 (15.2)	79.5 (12.8)	54.1 (16.7)
Concurrent task	ACC	93.1 (5.8)	93.3 (5.5)	89.0 (6.0)	87.8 (7.7)	89.0 (6.2)	88.4 (7.9)	76.1 (12.2)	74.1 (12.6)
	RT	586 (81.2)	610 (86.3)	664 (75.1)	681 (77)	652 (71.7)	672 (66.3)	743 (74.7)	748 (67.6)
Recall latency	Latency 1	1965 (485)	2043 (611)	1951 (413)	2270 (719)	4616 (3204)	4904 (3315)	4030 (1698)	6502 (5845)
	Latency 2	1052 (371)	1241 (554)	999 (387)	1306 (559)	1073 (549)	2948 (1853)	1092 (695)	2285 (2369)
	Latency 3	855 (408)	1655 (998)	1953 (356)	1883(1385)	1710 (2177)	2642 (3343)	2012 (2285)	2786 (3136)
	Latency 4	1718 (1278)	2396 (2489)	1666 (1280)	2281 (1928)	2412 (3735)	3666 (5389)	2198 (3402)	3963 (5385)
Type of errors	Omission	5.4 (6.4)	18.0 (10.1)	5.9 (6.2)	19.9 (11.3)	15.4 (9.6)	36.5 (15.6)	17.9 (11.7)	39.8 (17.8)
	Transposition	17.4 (9.8)	18.4 (8.8)	22.9 (14.3)	22.6 (13.1)	21.5 (12.7)	17.2 (11.2)	24.5 (14.2)	18.8 (10.7)
	Semantic	3.1 (3.9)	1.8 (2.4)	2.1 (2.2)	1.3 (1.6)	1.8 (2.8)	1.8 (2.7)	2.7 (4.4)	1.4 (2.0)
	Phonological	0 (0)	0.3 (0.8)	0 (0)	0.4 (0.9)	0 (0)	0.2 (0.6)	0.3 (0.9)	0.3 (1.0)
	Intrusion	0.6 (1.6)	4.1 (4.3)	0.4 (0.9)	3.7 (4.7)	0.3 (1.1)	2.9 (4.4)	0.6 (1.6)	2.4 (2.8)

Note. Recall performance is expressed as span and as percentage of correct serial and of free recall. The concurrent task performance is expressed as percentage of correct responses (ACC). RT stands for reaction time in milliseconds in the concurrent task. Omission, transposition, semantic, phonological, and intrusion refer to the percentage of recall errors for each type of errors. Standard deviations are in brackets.

Table 3. $BF_{inclusions}$ of all variables of interest for recall performance (span, percentage of serial and free recall), recall latency, recall error, percentage of accuracy (ACC) and reaction time (RTs in ms) in the concurrent tasks in Experiment 1.

Effect	Span	Serial recall	Free recall	Recall latency	Recall error	ACC	RTs
Concurrent task	43.3	313	1.97	0.15	7.98	1.45×10^{25}	3.34×10^{38}
Relatedness	1.26×10^{31}	2.10×10^{37}	1.63×10^{71}	1.90	1.22×10^9	0.36	69.7
Age	195	79.0	4447	1.09	78:4	2477	116
Concurrent task x Relatedness	0.27	0.12	0.26	0.19	0.13	0.32	0.42
Concurrent task x Age	0.19	0.09	2.87	0.20	0.16	1.34×10^8	0.29
Relatedness x Age	0.21	0.13	93.5	0.28	0.21	0.23	0.25
Concurrent task x Relatedness x Age	0.46	0.12	0.23	0.39	0.18	0.24	0.28
List length	-	3.93×10^{88}	4.69×10^{12}	-	-	-	-
List length x Concurrent task	-	0.004	0.008	-	-	-	-
List length x Relatedness	-	0.62	4311	-	-	-	-
List length x Age	-	0.003	1.37	-	-	-	-
List length x Concurrent task x Relatedness	-	0.14	0.04	-	-	-	-
List length x Concurrent task x Age	-	0.01	0.03	-	-	-	-
List length x Relatedness x Age	-	0.02	0.03	-	-	-	-
List length x Concurrent task x Relatedness x Age	-	0.03	0.14	-	-	-	-
Serial position	-	-	-	1.57×10^7	-	-	-
Serial position x Concurrent task	-	-	-	0.06	-	-	-
Serial position x Relatedness	-	-	-	0.16	-	-	-
Serial position x Age	-	-	-	4.11×10^5	-	-	-
Serial position x Concurrent task x Relatedness	-	-	-	0.12	-	-	-
Serial position x Concurrent task x Age	-	-	-	0.11	-	-	-
Serial position x Relatedness x Age	-	-	-	1.26	-	-	-
Serial position x Concurrent task x Relatedness x Age	-	-	-	0.19	-	-	-
Types of error	-	-	-	-	0.05	-	-
Types of error x Concurrent task	-	-	-	-	0.18	-	-
Types of error x Relatedness	-	-	-	-	1.45×10^{20}	-	-
Types of error x Age	-	-	-	-	2.22×10^{11}	-	-
Types of error x Concurrent task x Relatedness	-	-	-	-	0.23	-	-
Types of error x Concurrent task x Age	-	-	-	-	0.41	-	-
Types of error x Relatedness x Age	-	-	-	-	77:30	-	-
Types of error x Concurrent task x Relatedness x Age	-	-	-	-	0.29	-	-

Table 4. Descriptive statistics of all dependent variables (DVs) in Experiment 2 as a function of age group (adults vs. children), pace of the concurrent task (slow vs. fast) and associative relatedness of word lists (related vs. unrelated words).

DVs		Adults				Children			
		Slow Pace		Fast Pace		Slow Pace		Fast Pace	
		Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
Recall	Serial recall	82.3 (10.3)	71.5 (11.2)	70.1 (13.5)	58.8 (16.5)	38.9 (13.9)	23.0 (11.9)	31.8 (12.6)	17.0 (12.6)
	Free recall	91.6 (6.7)	77.5 (9.6)	85.0 (8.7)	66.1 (12.9)	59.2 (15.2)	31.4 (14.0)	57.7 (15.1)	28.8 (13.0)
Concurrent task	ACC	95.5 (3.3)	94.5 (3.6)	89.4 (5.8)	87.5 (7.3)	74.9 (14.1)	74.4 (15.1)	59.9 (13.7)	59.5 (13.8)
	RT	624 (98.2)	624 (98.3)	534 (45.3)	541 (47.7)	769 (128.0)	764 (131.0)	543 (72.7)	532 (72.0)
Recall latency	Latency 1	3232 (2234)	3094 (1656)	3732 (2682)	3761 (2893)	2808 (1188)	3827 (1615)	3058 (1643)	3599 (1689)
	Latency 2	1633 (839)	2100 (1310)	2003 (1149)	2087 (948)	2973 (1476)	3970 (2696)	3179 (1533)	4154 (1581)
	Latency 3	2000 (950)	2818 (2785)	2193 (1224)	3378 (3518)	2971 (1313)	4828 (3248)	2884 (2043)	4159 (2710)
	Latency 4	1481 (959)	1908 (1059)	1696 (1059)	2152 (1519)	2711 (1802)	3422 (1952)	2990 (2060)	2703 (1504)
Type of errors	Omission	6.1 (6.3)	20.1 (9.9)	12.2 (7.6)	32.3 (14.2)	35.6 (15.3)	61.0 (14.1)	37.2 (16.4)	64.0 (13.7)
	Transposition	9.1 (7.4)	5.1 (4.9)	14.7 (7.7)	7.5 (6.7)	19.4 (18.3)	9.5 (11.0)	24.7 (17.6)	12.1 (13.4)
	Semantic	2.2 (2.4)	1.0 (1.7)	2.8 (3.0)	0.4 (1.1)	4.1 (3.7)	1.1 (1.5)	3.9 (3.9)	1.1 (2.0)
	Phonological	0 (0)	0.7 (1.5)	0 (0)	0 (0)	0.7 (1.6)	2.2 (2.9)	0.5 (1.2)	1.8 (3.0)
	Intrusion	0.3 (1.2)	1.0 (2.6)	0.1 (0.6)	0.5 (1.4)	0.3 (0.9)	1.0 (2.7)	1.22 (1.8)	2.0 (2.9)

Note. Recall performance is expressed as percentage of correct serial and of free recall. The concurrent task performance is expressed as

percentage of correct responses (ACC). RT stands for reaction time in milliseconds in the concurrent task. Omission, transposition, semantic,

phonological, and intrusion refer to the percentage of recall errors for each type of errors. Standard deviations are in brackets.

Table 5. $BF_{inclusions}$ of all variables of interest for recall performance (span, percentage of serial and free recall), recall latency, recall error, percentage of accuracy (ACC) and reaction time (RT) in concurrent task in Experiment 2.

Effect	Serial recall	Free recall	Recall latency	Recall error	ACC	RT
Pace	4.47×10^5	3295	0.15	4.25×10^3	1.54×10^{33}	3.32×10^{40}
Relatedness	2.43×10^{17}	5.13×10^{41}	2.90×10^5	1.43×10^4	0.44	0.12
Age	1.62×10^{21}	5.04×10^{18}	118	4.26×10^{14}	1.37×10^{11}	127
Pace x Relatedness	0.45	0.41	0.19	0.17	0.21	0.18
Pace x Age	123	14.4	0.56	0.53	8.76×10^7	4.39×10^{16}
Relatedness x Age	1.94	3.76×10^4	0.89	0.21	0.27	0.31
Pace x Relatedness x Age	0.11	0.35	0.13	0.24	0.28	0.30
Serial position	-	-	2.19×10^6	-	-	-
Serial position x Pace	-	-	0.001	-	-	-
Serial position x Relatedness	-	-	1.55	-	-	-
Serial position x Age	-	-	1.20×10^4	-	-	-
Serial position x Pace x Relatedness	-	-	0.02	-	-	-
Serial position x Pace x Age	-	-	0.04	-	-	-
Serial position x Relatedness x Age	-	-	0.08	-	-	-
Serial position x Pace x Relatedness x Age	-	-	0.07	-	-	-
Types of error	-	-	-	5.23×10^{30}	-	-
Types of error x Pace	-	-	-	0.24	-	-
Types of error x Relatedness	-	-	-	3.34×10^{30}	-	-
Types of error x Age	-	-	-	1.07×10^{22}	-	-
Types of error x Pace x Relatedness	-	-	-	0.64	-	-
Types of error x Pace x Age	-	-	-	0.59	-	-
Types of error x Relatedness x Age	-	-	-	27.41	-	-
Types of error x Pace x Relatedness x Age	-	-	-	0.43	-	-

