

Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory

Valérie Camos · Gerome Mora · Klaus Oberauer

© The Psychonomic Society 2010

Abstract Because both articulatory rehearsal and attentional refreshing aid in the maintenance of verbal information in the short term, the present study evaluated the adaptive use of these mechanisms, using a complex span paradigm. In [Experiment 1](#), the phonological similarity of memory list words and the attentional demand of concurrent processing were manipulated. As was predicted, a phonological similarity effect (PSE) appeared only when the concurrent task was attention demanding, thus impairing the use of refreshing and encouraging rehearsal. To verify that PSE indicates the use of rehearsal, participants were instructed to use one of the two mechanisms in [Experiments 2 and 3](#). In accordance with [Experiment 1](#), the PSE was observed only under rehearsal. Thus, adults could adaptively choose between the two mechanisms. When remembering phonologically confusable materials, they prefer refreshing in order to reduce the impact of phonological characteristics. When available attention is

reduced, they favor a less attention-demanding mechanism, rehearsal.

Keywords Working memory · Rehearsal · Refreshing · Phonological similarity · Strategy

Working memory is a system dedicated to the maintenance of information in the context of concurrent processing. Within this system, two mechanisms have been described as responsible for the maintenance of verbal information. Some models have proposed a mechanism specialized for the verbal domain (articulatory rehearsal; Baddeley, 1986), whereas others have introduced a general attention-based mechanism (attentional refreshing; Barrouillet & Camos, 2007; Cowan, 1999, 2005; Johnson, 1992). Recently, these two mechanisms were shown to be independent in verbal working memory tasks (Camos, Lagner, & Barrouillet, 2009; Hudjetz & Oberauer, 2007). The aim of the present study was to test the hypothesis that adults could adaptively use either of these two mechanisms, according to the characteristics of the tasks or according to instructions.

V. Camos · G. Mora
Université de Bourgogne,
Dijon, France

V. Camos
Institut Universitaire de France,
Paris, France

K. Oberauer
University of Bristol,
Bristol, England

V. Camos (✉)
Université de Bourgogne, LEAD-CNRS, Pôle AAFE,
Esplanade Erasme,
B.P. 26513, 21065 Dijon Cedex, France
e-mail: valerie.camos@u-bourgogne.fr

Articulatory rehearsal and the phonological similarity effect

In the short-term memory literature, there is a long tradition of assuming that memory for verbal information relies on subvocal rehearsal (Atkinson & Shiffrin, 1968; Sperling, 1967; Waugh & Norman, 1965). In the classic model of working memory (Baddeley, 1986), the maintenance of verbal information relies on a similar mechanism, articulatory Rehearsal, that maintains memory traces in the store of

the phonological loop. The concept of the phonological loop attracted considerable research during the past 30 years. One piece of evidence for its phonological nature came from the phonological similarity effect (PSE; although see Macken & Jones, 2003). First, Conrad and Hull (1964) observed, in a short-term memory task, that consonants similar in sound are more difficult to maintain than letters that sound different (see also Conrad, 1964). Baddeley (1966b) found that lists of phonologically similar words (e.g., *mad, man, mat, cap, cad, can, cat*) are recalled more poorly than lists of dissimilar words (e.g., *cow, day, bar, few, hot, pen, soup, pit*). There is meanwhile a large body of literature reporting this PSE with so-called *simple span* tasks (i.e., the immediate serial recall of lists). In contrast, studies investigating this effect with *complex span* tasks, which combine list recall with a concurrent-processing task to tap working memory (e.g., reading span, listening span, operation span), are scarce. Recently, a PSE was observed by Loble, Baddeley, and Gathercole (2005) in a listening span task and by Copeland and Radvansky (2001) in an operation span task. However, Copeland and Radvansky failed to find such an effect in a reading span task. Likewise, Tehan, Hendry, and Kocinski (2001) found no PSE in an operation span task.

There are two families of theoretical explanations for the PSE, one attributing it to effects at encoding and the other attributing it to effects at retrieval. In Baddeley's (1986) model of the phonological loop, the PSE has been explained by assuming trace discrimination problems within the store of the phonological loop that arise during retrieval (Gathercole & Baddeley, 1993; Salamé & Baddeley, 1982). More recent models attribute the effect of similarity, at least in part, to processes during encoding. For instance, in the serial-order-in-a-box (SOB) model (Farrell & Lewandowsky, 2002) encoding is novelty-gated, so that items that are similar to already encoded items are encoded with reduced strength (Farrell & Lewandowsky, 2003). In feature-overwriting models (Nairne, 1990; Oberauer & Kliegl, 2006), similar items share more features with each other than do dissimilar items, resulting in more mutual degradation of representations through feature overwriting.

Regardless of which explanation is correct, a PSE can arise only if memory performance relies to a substantial degree on phonological representations. Because the PSE is also found with visually presented verbal materials, some mechanism for generating phonological representations from the orthographic input must be assumed. In Baddeley's (1986) model, the same mechanism, subvocal articulation, is responsible for phonological recoding of visual input and for rehearsal. A second prerequisite for a phonological similarity effect is that phonological representations, once generated during encoding, are actually used for retrieval. Articulatory

rehearsal is a likely mechanism for protecting phonological memory traces against forgetting. Therefore, measures to prevent articulatory rehearsal, such as concurrent articulatory suppression, should reduce the PSE, because it forces people to rely on other, nonphonological memory representations. In confirmation of this prediction, the PSE disappears when a verbal-shadowing task is introduced into the retention interval in a delayed recall task (Fournet, Juphard, Monnier, & Roulin, 2003; Tehan & Humphreys, 1998) or as the processing component of a complex span task (Mora, Barrouillet, & Camos, 2009). However, contrary to these findings, Fallon, Groves, and Tehan (1999) reported a PSE in a delayed recall task in which the retention interval was filled with a digit-reading task, and Fournet et al. still found a phonological similarity effect with their shorter retention delay. To conclude, studies on the PSE in working memory are sparse and do not lead to a consistent picture. This discrepancy in the findings might arise because people avail themselves of a variety of maintenance mechanisms and memory representations on which they can rely to maintain verbal information in working memory.

An attention-based mechanism of maintenance

Indeed, besides articulatory rehearsal, maintenance of verbal material can also benefit from another mechanism, referred to as *attentional refreshing* (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Barrouillet & Camos, 2001; Cowan, 1999, 2005; Johnson, 1992; Johnson et al., 2005; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Raye, Johnson, Mitchell, Reeder, & Greene, 2002). The common idea of the different theories cited above is that attention serves to maintain memory traces in an active state, because directing attention briefly to a target strengthens the target's representation in working memory. According to Cowan, when the trace of an item is activated in memory, its activation is maintained as long as the item is within the focus of attention. As soon as the focus of attention is switched away, the trace begins to fade away until it is completely lost. Before complete loss, the trace could be reactivated by focusing attention on it again. Similarly, in Johnson's model, refreshing is conceived of as thinking briefly of an item, thereby increasing its level of activation. The time-based resource-sharing (TBRS) model proposed by Barrouillet et al. (2004; Barrouillet et al., 2007) endorsed this conception by assuming that refreshing through attentional focusing could reactivate the decaying memory traces. Furthermore, the TBRS model assumes that attention can be devoted to only one process at a time and that, because both maintenance and concurrent processing require attentional resources, attention is switched from one

to another in complex span tasks. As a consequence, recall in these tasks depends on the attentional demand introduced by the concurrent-processing task.

Conceptually, refreshing differs from rehearsal in three regards. First, refreshing requires a central attentional mechanism such that, when the cognitive system engages in refreshing, it cannot at the same time engage in another process that requires that attentional mechanism. Articulatory rehearsal, in contrast, demands no attention after a brief initial setup period (Naveh-Benjamin & Jonides, 1984). Second, articulatory rehearsal consists of sub vocal speech production and, thus, cannot proceed concurrently with overt speech (i.e., articulatory suppression, or reading unrelated material aloud), whereas attentional refreshing can proceed concurrently with unrelated overt speech (Hudjetz & Oberauer, 2007). Third, rehearsal can maintain only phonological representations, whereas refreshing can maintain any representations.

Empirically, the distinction between articulatory rehearsal and attentional refreshing gained direct support through neurophysiological and behavioral evidence. Brain-imaging studies showed that the two mechanisms are implemented in distinct brain areas. Whereas the ventrolateral prefrontal cortex (Brodmann's area [BA] 44 reflects a subvocal articulatory rehearsal of phonological information, the dorsolateral prefrontal cortex (BA 9) is assumed to reflect the involvement of attention in the maintenance of various activated information (Johnson et al., 2005; Romero, Walsh, & Papagno, 2006). Recent behavioral evidence reinforced this distinction (Camos et al., 2009; Hudjetz & Oberauer, 2007). Hudjetz and Oberauer, using a reading span task, and Camos et al., using computer-paced complex span tasks, manipulated separately the opportunity for rehearsal and for refreshing. These experiments never yielded any significant interaction between the varied factors, supporting the idea that articulatory rehearsal and attentional refreshing could be independently involved in the maintenance of verbal information. To account for this distinction between the two maintenance mechanisms, Camos et al. proposed that refreshing is a central and general-purpose mechanism, because it relies on attention and can be applied to any representational code, whereas rehearsal is peripheral, because it involves a subvocal articulatory process operating on a domain-specific code (i.e., phonological) and superficial levels of encoding. This extended version of the TBRS model could account for the fact that recall is impaired by variation of the attentional demand of any concurrent-processing task, even a nonverbal one (Barrouillet et al., 2004; Barrouillet et al., 2007; Camos et al., 2009), and that, at the same time, for findings of selective interference of verbal-processing tasks with verbal memory (reviewed in Baddeley, 1986; Vergauwe, Barrouillet, & Camos, 2010). This theory implies that the maintenance process applicable

to verbal information depends on the codes used to represent the information: All representations can be refreshed, whereas only phonological representations can be rehearsed through rearticulation.

Strategy effects on encoding and maintenance

Several studies have demonstrated the impact of encoding strategies on the PSE in immediate serial recall of visually presented items. For instance, Campoy and Baddeley (2008) found that phonologically dissimilar words were recalled better than similar words when participants were instructed to encode words phonologically. The induction of a semantic encoding strategy eliminated the PSE.

Similarly, Hanley and Bakopoulou (2003) observed a PSE only for an experimental group instructed in using a phonological strategy. Moreover, the group instructed to use the semantic strategy outperformed the participants using the phonological strategy. These authors concluded that, when instructed to do so, adults could adopt alternative strategies that abolish the PSE. Salamé and Baddeley (1986) suggested that when lists contain large numbers of phonologically similar items, a phonological strategy is likely to be difficult and will be abandoned in favor of a visual or semantic strategy, explaining the disappearance of the PSE at long list lengths. In line with this suggestion, Baddeley (1966a) found that with long sequences, serial recall is affected by semantic, rather than phonological, similarity between items. To summarize, in immediate serial recall tasks, the encoding strategy, phonological versus semantic, varies according to either the difficulty of the task (short vs. long lists) or the instructions.

Therefore, we suggest that in complex span tasks, which are typically more difficult than simple span tasks of similar list length, participants will tend to avoid relying on phonological codes as much as possible, especially if the material to be maintained is phonologically similar. The chosen representational format (phonological or non phonological) has consequences for the available maintenance mechanism: Nonphonological codes can be maintained only through attentional refreshing, whereas phonological codes can also be maintained through rehearsal. Therefore, the choice of representational format in a complex span task is likely to also depend on the opportunity for attentional refreshing and for articulatory rehearsal. When attention is available for refreshing, nonphonological representations, together with attentional refreshing, should be favored, instead of phonological representations together with articulatory rehearsal. As a consequence, little or no PSE should be observed in a complex span task when the attentional demand of the processing component is

modest. When the attentional demand is high, in contrast, the opportunity for refreshing is reduced, and this renders reliance on the alternative maintenance mechanism, rehearsal, more attractive, especially since articulatory rehearsal is known to be a low-demanding process; only the retrieval of the phonological traces in long-term memory requires some attention (Baddeley, 2007; Murray, 1968; Naveh-Benjamin & Jonides, 1984). Because rehearsal requires—and generates—a phonological representation, a PSE would be the consequence of that strategy choice. Therefore, we predict that the PSE in a complex span task is larger when the attentional demand of the processing component is higher. In addition, we assume that instructions can modulate people's choice of a representation format and an associated maintenance strategy. When people are instructed to rely on articulatory rehearsal, as in Experiment 2, they will abandon their preference for a nonphonological representation and, instead, will rely on phonological codes that can be rehearsed. Therefore, we predicted that with instructions to use phonological rehearsal, a PSE would emerge, regardless of the attentional demand of the processing component of a complex span task. On the contrary, when people were instructed to rely on attentional refreshing, as in Experiment 3, they would rely on nonphonological representations that they could refresh. The PSE should then disappear, despite the attentional demand of the concurrent processing.

Experiment 1

To test that the adaptive choice between articulatory rehearsal and attentional refreshing depends on the attentional demand of the concurrent task, we used a complex span paradigm in which the processing component was either a choice reaction time (CRT) task or an attentionally less demanding simple reaction time (SRT) task and the material to be maintained consisted of lists of phonologically similar or dissimilar words. Participants did not receive any instruction on the maintenance mechanism to use, but the variation of the concurrent attentional demand should vary the opportunity for refreshing. We predicted that with a low concurrent attentional demand (i.e., with the SRT task), people would have the opportunity to engage in refreshing and would prefer relying on refreshing of nonphonological representations in order to circumvent the difficulty arising from the confusability of phonologically similar items. In contrast, with high concurrent attentional demand (i.e., with the CRT task), the opportunity for refreshing of nonphonological representations would be much reduced. As a consequence, people would have to rely on articulatory rehearsal, which boosts only the phonological representation of memory items. As a

consequence, the PSE should be larger with the CRT than with the SRT task.

We manipulated the phonological characteristics of the list items in two ways. First, phonological similarity was operationalized as phonological neighborhood between list items. Phonological neighbors are words that differ in only one phoneme. We created similar lists by maximizing the average number of neighbors that each list item had among the other items on the same list. Second, we manipulated phoneme overlap—that is, the number of phonemes of each list word that occurred again in any other list word. Whereas high pairwise similarity (defined through neighborhood) implies high phoneme overlap, the opposite is not the case, because high phoneme overlap can be created by repeating phonemes of one word in a distributed fashion in other list words and in different within-word positions. We were interested in these two kinds of phonological characteristics because they could speak to the mechanism of interference between phonological representations. If interference acts through confusion of whole words at retrieval, phonological neighborhood should be the relevant factor. If interference acts through feature overwriting during encoding (Nairne, 1990; Oberauer & Kliegl, 2006), high phoneme overlap should create interference even for low-confusable items. Evidence for feature overwriting has been obtained primarily with paradigms investigating interference between memory items and distractors (Lange & Oberauer, 2005; Oberauer, 2009; Oberauer & Lange, 2008); here, we explored for the first time the effect of manipulating phoneme overlap between all items of a memory list. Thus, three types of lists of words were created: lists with many phonological neighbors (high similarity, HS), lists with words that were not neighbors but shared many phonemes (high overlap, HO), and lists with words that were not neighbors and shared few phonemes (low overlap, LO).

Method

Participants

Twenty undergraduate students at the University of Bristol (17 women, 3 men) received partial course credit or money for participating. They were all English native speakers, between 18 and 24 years of age ($M = 20.1$, $SD = 1.52$).

Materials

Lists of six words were constructed from a pool of 1,124 monosyllabic singular English nouns (CELEX), excluding

words with strong emotional meaning. Twenty-four lists were constructed for each type of lists (HS, HO, and LO). As in Oberauer (2009) and Oberauer and Lange (2008), the same words were used to build the three different types of lists but were arranged in the lists in different ways (Table 1). Therefore, across all lists, there were no differences due to word characteristics such as length, concreteness, or imageability. Every word in the HS lists had an average 2.11 of phonological neighbors (range: 2–2.67), and these lists had an average of 8.94 phoneme overlaps (range: 5.33–13.67). In contrast, words in the HO and LO lists had, on average, less than one phonological neighbor in the list (range: 0–1). Lists in the HO condition had an average of 4.96 phoneme overlaps (range: 2.33–7.33), whereas lists in the LO condition had an average of 1.80 phoneme overlaps (range: 0–6.33). One half of the participants were presented with 12 lists from each condition, and the other half with the remaining 12 lists. For each participant in each condition, 6 lists were presented with a CRT task and 6 lists with an SRT task. The order of presentation of CRT and SRT tasks were randomized, with the constraint that the same reaction time (RT) task could not be performed more than twice in a row. The order of presentation of the different lists was also randomized within the CRT and SRT conditions. Finally, within a list, the words were displayed in a random order. Due to the repeated use of the same set of words for the three conditions, a word could be seen more than once throughout the experiment (on average, 1.94 times per participant).

Procedure

Participants were seated about 60 cm from a computer screen and were presented with the tasks using PsyScope software (Cohen, MacWhinney, Flatt, & Proost, 1993). Each series began with an asterisk centered in the screen for 500 ms, followed by the first word of a list, presented in red for 1,500 ms. After a postword delay of 333 ms, six stimuli appeared one by one for 666 ms each, followed by a 334-ms delay, for a total duration of 1,000 ms per stimulus. Such series of six stimuli appeared after each word. These stimuli consisted of 20×20 mm black squares either centered on the screen in the SRT task or centered on one of two possible locations in the upper or the lower part of the screen (15 mm apart from the center of the screen) in the CRT task. In each series of CRT trials, the squares appeared in the two locations with the same frequency. After 6 squares, the following word was displayed for 1,500 ms, and so on. The interword interval was 6,333 ms ($333 + 6 \times 1,000$). At the end of a series, participants were instructed to recall words in the same order as that in which they had been presented. Thus,

“1” appeared on the screen, and participants had to typewrite the first word to be remembered, using a keyboard. When they finished typewriting the first word, they had to press “Enter” to go to the second word, “2” appeared, and so on. If participants were not able to remember a word, they pressed “Enter” to go to the next one. They were informed that they could not go back to previous words after having pressed “Enter.” After having typed the last word, participants pressed the space bar to start the next series. At the beginning of each series, participants were informed about the task to perform (e.g., “one finger” for the SRT task, “two fingers” for the CRT task). Participants were instructed to memorize each word and to press a key each time the square appeared in the SRT task or to judge the location of each square as accurately as possible by pressing either a left- or right-hand key for the upper and the lower locations, respectively, in the CRT task. The RTs and accuracy were recorded.

A training phase familiarized participants with the RT tasks (one series with the SRT task, two series with the CRT task: one just watching the square locations and one pressing keys to respond) and then with the working memory span tasks (three series with each task, alternating SRT and CRT tasks). Words presented during training were not used in the test phase. The experiment lasted about 1 hr.

Results and discussion

Three participants were discarded from the analysis because they had fewer than 70% correct responses in the CRT task. The percentage of correct responses was higher in the SRT (98%) than in the CRT task (85%), $t(16) = 8.58$, $p < .001$. Because of a ceiling effect for the SRT task, we did not analyze its accuracy. However, a one-way analysis of variance (ANOVA) was performed on the percentage of correct responses for the CRT task. Percentages were not significantly different for the three types of lists (HS = 84%; HO = 86%; LO = 85%), $F(2, 32) = 1.15$, $p = .33$, $\eta^2 = .07$. A 2 (task) \times 3 (list type) ANOVA was performed on RTs for the correct responses. The effect of list type was not significant ($F < 1$), but the effect of task was. As was expected, RTs were longer for the CRT task than for the SRT task (429 vs. 309 ms), $F(1, 16) = 29.89$, $p < .001$, $\eta^2 = .99$, and the interaction between task and list was not significant, $F < 1$.

A 2 (task) \times 3 (list) ANOVA was performed on the rate of words recalled in the correct position. As we predicted, fewer words were recalled in the correct position for the CRT task than for the SRT task (68% and 82%, respectively), $F(1, 16) = 28.28$, $p < .001$, $\eta^2 = .70$, (Fig. 1). The effect of list type was significant, $F(2, 32) = 4.00$, $p = .03$, $\eta^2 = .17$. Recall was poorer for the HS lists (70%) than for the two other types of

Table 1 Lists of words used in Experiments 1 and 2 with the average number of neighbors and the average number of phoneme overlap

Condition	Neighbors	Overlap	Words						
High Similarity (HS)	2.00	7.33	ban	man	mass	gas	bag	match	
	2.00	9.67	twin	tin	town	tone	stone	phone	
	2.00	11.33	hind	hand	mind	land	wind	wine	
	2.00	10.33	sneer	sphere	spear	peer	year	pair	
	2.33	8.67	slide	side	size	sight	night	light	
	2.00	6.33	coal	role	road	rule	room	hole	
	2.00	13.33	bream	dream	cream	crime	creek	crack	
	2.33	7.00	thumb	sum	son	sun	scene	sea	
	2.00	6.00	charm	arm	art	part	heart	heat	
	2.00	6.67	toe	tea	team	time	type	term	
	2.00	7.00	hawk	hall	wall	war	door	ball	
	2.00	5.67	meal	male	meat	seat	suit	beat	
	2.00	9.33	corpse	course	court	horse	house	source	
	2.33	13.67	draught	craft	draft	raft	shaft	grub	
	2.00	7.33	date	day	gate	game	name	gain	
	2.00	5.33	hay	bay	boy	base	face	race	
	2.00	7.00	mole	soul	soil	oil	cell	bell	
	2.00	10.67	cramp	camp	cap	cup	cash	cat	
	2.67	13.00	hump	jump	pump	lump	lamp	lamb	
	2.00	9.33	drug	rug	drum	jug	judge	mug	
	2.33	8.33	phrase	phase	faith	fate	foot	fight	
	2.00	9.33	toil	tail	sale	scale	school	skill	
	2.67	11.33	root	fruit	roof	route	proof	boot	
	2.00	10.67	bunk	bank	rank	tank	tank	rack	
	Mean	2.11	8.94						
	High Overlap (HO)	0.00	5.00	fight	ball	stone	male	side	term
0.00		6.00	sea	course	wall	phone	hawk	camp	
0.00		5.33	seat	thumb	town	source	sneer	rack	
0.00		6.00	gain	bream	charm	crime	art	slide	
0.00		5.00	twin	light	cream	rule	hall	meal	
0.00		6.00	cash	cup	hand	wind	hind	night	
0.00		6.67	date	tin	room	crack	creek	wine	
0.00		5.00	sight	tone	name	time	meat	tea	
0.00		7.33	bulk	day	sun	door	root	drug	
0.00		5.33	bank	gas	craft	shaft	boot	cell	
0.00		5.33	mug	part	heat	toe	type	team	
0.00		5.33	soil	sphere	dream	draft	face	phrase	
0.00		5.67	year	role	bag	grub	draught	rug	
0.00		5.00	boy	horse	base	land	oil	bunk	
0.00		5.00	rank	heart	house	beat	gate	cat	
0.00		4.00	toil	hole	bay	fate	scale	route	
0.00		4.00	phase	mass	peer	sale	school	cramp	
0.00		5.67	tail	size	coal	game	mole	corpse	
0.00		4.33	bell	match	hump	sum	drum	tank	
0.00		4.33	court	scene	jump	lamp	ban	proof	
0.00		4.00	judge	spear	pair	war	pump	mind	

Table 1 (continued)

Condition	Neighbors	Overlap	Words					
	0.00	3.33	skill	jug	suit	soul	race	fruit
	0.00	3.00	hay	road	raft	faith	roof	foot
	0.00	2.33	son	man	lamb	lump	cap	arm
Mean	0.00	4.96						
Low Overlap (LO)	0.00	0.00	seat	match	date	role	bunk	rack
	0.00	0.00	roof	man	mole	slide	lamp	lump
	0.00	0.00	dream	gas	twin	soul	thumb	foot
	0.00	0.00	scene	bag	pair	light	road	charm
	0.00	0.00	court	crack	tin	hole	art	jug
	0.00	0.00	mass	town	rule	door	boy	cup
	0.00	0.00	sight	phone	peer	hall	cap	boot
	0.00	1.00	hump	tone	sphere	creek	wall	bay
	0.00	0.00	face	hind	bream	part	oil	cash
	0.00	0.00	time	hand	sea	ball	rug	phase
	0.00	0.67	course	mind	heart	grub	faith	tank
	0.00	0.00	heat	land	size	arm	bulk	proof
	0.00	1.67	craft	ban	room	toe	judge	rank
	0.00	1.33	sale	wind	cream	tea	bell	draft
	0.00	1.00	wine	sneer	meal	drug	coal	bank
	1.00	4.00	soil	stone	game	year	fruit	root
	1.00	5.33	type	team	gain	route	scale	tail
	0.33	3.00	school	night	camp	skill	war	day
	0.00	2.00	phrase	name	beat	fate	mug	drum
	0.67	4.67	source	side	spear	toil	fight	base
	0.33	5.00	gate	crime	sum	jump	cat	cramp
	0.67	4.00	suit	term	pump	lamb	race	hawk
	0.67	6.33	sun	male	cell	shaft	raft	hay
	0.00	3.33	house	son	draught	horse	corpse	meat
Mean	0.19	1.81						

Note: "Neighbors" is the average number of phonological neighbors in a list. "Overlap" is the average number of phoneme overlaps each word has with other words on the list

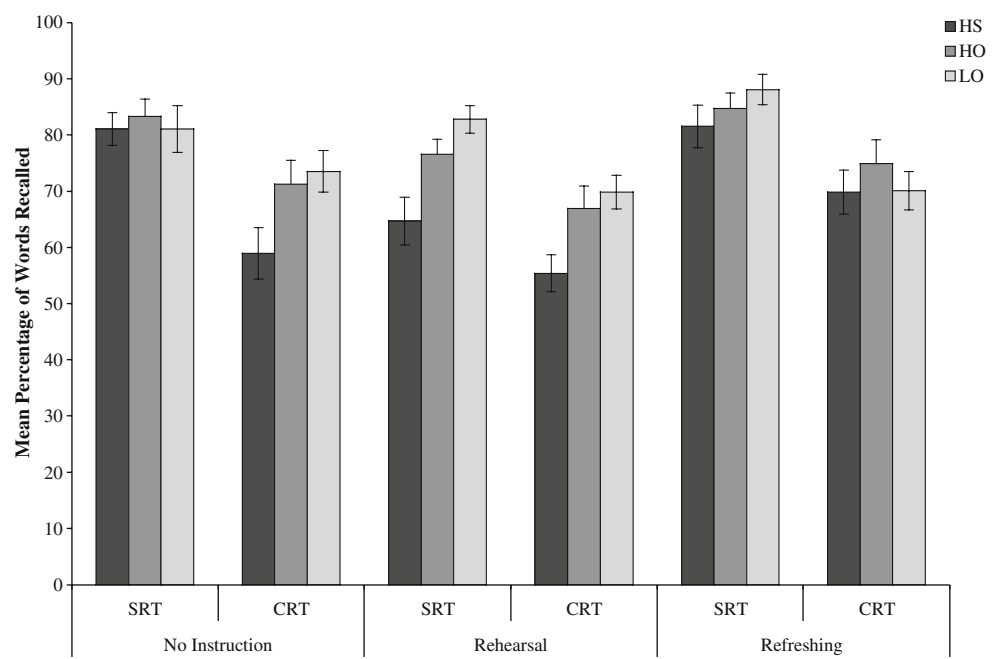
lists (HO, 77%; LO, 77%), which did not differ, $F(1, 16) = 19.39$, $p < .001$, and $p = 1$, respectively, $\eta^2 = .17$, and $p = 1$. Finally, the interaction between task and list was significant, $F(2, 32) = 3.42$, $p < .05$, $\eta^2 = .13$. The effect of list was significant on the trials with the CRT task, $F(2, 32) = 6.18$, $p < .01$, $\eta^2 = .29$, but not on those with SRT task, $F < 1$. More specifically, on trials with the CRT task, the list type effect arose from the significant difference between the HS lists and the two other types of lists, $F(1, 16) = 15.12$, $p < .01$, $\eta^2 = .29$, the difference between HO and LO lists being nonsignificant, $F < 1$.

The lack of a PSE was not due to a trade-off between item and order information, since it was observed for rhyming words (e.g., Fallon et al., 1999). The pattern of findings was similar for both item score (i.e., recall regardless position) and order accuracy (i.e., dividing the correct-in-position score by the item score) as for recall in

the correct position. Item scores were poorer for the HS lists (79%) than for the two other types of lists (HO = 83% and LO = 83%), $F(1, 16) = 13.49$, $p < .01$, $\eta^2 = .13$. Likewise, order accuracy was poorer for the HS lists (88%) than for the two other types of lists (HO = 93% and LO = 92%), $F(1, 16) = 6.83$, $p < .05$, $\eta^2 = .25$. No difference was apparent between the HO and LO lists on either method of scoring, $F_s < 1$.

Moreover, the effect of list was significant for the trials in the CRT task for the item score (HS = 71%, HO = 78%, and LO, 81%), $F(2, 32) = 4.78$, $p < 0.5$, $\eta^2 = .25$, and for the order score (HS = 82%, HO = 91%, LO = 90%), $F(2, 32) = 4.09$, $p < .05$, $\eta^2 = .48$, but not in those in the SRT task, $F_s < 1$, for item scores (HS = 86%, HO = 88%, and LO = 86%) and order accuracy (HS = 94%, HO = 94%, LO = 93%), respectively. More specifically, for the trials in the CRT task, the list-type effect arose from the

Fig. 1 Mean percentages of words recalled in the correct position according to the task (serial reaction time [SRT] vs. choice reaction time [CRT]) and the lists (HS = high-similarity lists, HO = low-similarity lists with high overlap, and LO = low-similarity lists with low overlap) when participants received no instructions (Experiment 1) and when they were instructed to use rehearsal (Experiment 2) or refreshing (Experiment 3). Bars represent standard errors



significant difference between HS lists and the two other types of lists, $F(1, 16) = 12.84$, $p < .01$, $\eta^2 = .23$, for item scores, and $F(1, 16) = 5.84$, $p < .05$, $\eta^2 = .48$, for order memory. The difference between HO and LO lists was nonsignificant, $F_s < 1$, for both item and order scores.

Thus, as we predicted, the phonological characteristics of the materials to be maintained affected recall, but only when attention was required to perform a concurrent task. In contrast, when attention was available for maintenance, the level of recall remained similar regardless of the phonological similarity of the words in the lists. Such a disappearance of the PSE when the concurrent task is attention demanding was already observed by Tehan et al. (2001, Experiment 2). These findings are in agreement with our hypothesis that adults can adaptively use rehearsal or refreshing to maintain verbal information in the short term. When the materials to be maintained can lead to phonological confusion, as it could here, refreshing is favored, because it can be applied to nonphonological representations. However, if attention is not available for refreshing, participants could back up to rehearsal, a less attention-demanding mechanism.

In sum, our first experiment provided evidence supporting the assumption that people adaptively shift their preferences between rehearsal and refreshing in response to the demands of the task. In the remaining two experiments, we directly tested the assumption that people's preference for one of the two maintenance mechanisms is under their intentional control, rather than being caused directly by task properties. In Experiment 2, we instructed participants to use articulatory rehearsal, and in Experiment 3, we instructed them to use refreshing. We predicted that a

phonological similarity effect should be found in Experiment 2, regardless of task demands, and that it should be absent in Experiment 3, regardless of task demands.

Experiment 2

In the second experiment, we explicitly asked participants to use articulatory rehearsal for maintaining the lists of words—that is, to subvocally repeat the words. If the task \times list interaction observed in Experiment 1 relied on an adaptive choice between refreshing and articulatory rehearsal, this interaction should disappear with this instruction, and recall should be reduced for lists in which words are phonologically similar even when sufficient attention is available.

Method

Participants

Twenty-three undergraduate students at the University of Bristol (18 women, 5 men) received partial course credit or money for participating. They were all English native speakers, between 18 and 29 years of age ($M = 19.91$, $SD = 2.41$). None of them had participated in Experiment 1.

Materials and procedure

The materials and procedure were the same as those in Experiment 1, except that participants were instructed to

use subvocal rehearsal to maintain the words by constantly repeating these words silently in their heads.

Results and discussion

Three participants were discarded from the analysis because they had fewer than 70% correct responses in the CRT task. The pattern of results for the percentages of correct responses and for the RTs for the RT tasks was similar to that in [Experiment 1](#). The percentage of correct responses was higher in the SRT task (97%) than in the CRT task (82%), $t(19) = 10.94$, $p < .001$. The one-way ANOVA performed on percentage of correct responses for the CRT task revealed that the three types of lists did not differ (HS = 83%, HO = 83%, and LO = 80%), $F(2, 38) = 2.16$, $p = .13$, $\eta^2 = .10$. Similarly, for the RTs, the effect of list was not significant, $F < 1$, but the effect of task was significant. The RTs were longer in the CRT task than in the SRT task (438 vs. 330 ms), $F(1, 19) = 31.79$, $p < .001$, $\eta^2 = .99$. The interaction between task and list was not significant, $p = .27$.

As in [Experiment 1](#), a 2 (task) \times 3 (list type) ANOVA was performed on the rate of words recalled in the correct position. Recall performance was poorer for trials in the CRT task than for trials in the SRT task (64% and 75%, respectively), $F(1, 19) = 27.36$, $p < .001$, $\eta^2 = .37$, (Fig. 1). As in [Experiment 1](#), the overall effect of list type was significant, $F(2, 38) = 31.99$, $p < .001$, $\eta^2 = .62$. Recall performance was poorer for the HS lists (60%) than for the two other types of lists (74%), $F(1, 19) = 40.66$, $p < .001$, $\eta^2 = .57$, and poorer for the HO lists (72%) than for the LO lists (76%), $F(1, 19) = 8.80$, $p < .01$, $\eta^2 = .05$. As was predicted from the assumption of adaptive choice of the maintenance process, the list type effect did not significantly differ between the two tasks, $F < 1$.

As in [Experiment 1](#), the PSE was the same in the item and order scores as in the correct-in-position score. Item scores were poorer for HS lists (73%) than for the other two types of lists, (HO : 80% and LO : 83%), $F(1, 19) = 16.81$, $p < .001$, $\eta^2 = .42$, $p < .$. The difference between the HO and LO lists was also significant, $F(1, 19) = 6.05$, $p < .05$, $\eta^2 = .06$. Likewise, order scores were poorer for the HS lists (81%) than for two other types of lists (HO = 90% and LO = 92%), $F(1, 19) = 22.59$, $p < .001$, $\eta^2 = .78$, and order recall was slightly better for the LO lists than for the HO lists, $F(1, 19) = 4.61$, $p < .05$, $\eta^2 = .03$. The interaction between task and type of lists was not significant in either the item or the order scores, $F_s < 1$.

Contrary to [Experiment 1](#), the phonological characteristics of the lists did not interact with the attentional demands of the concurrent task. For both the SRT and CRT tasks, the recall was poorer for lists of phonologically

similar words than for lists of dissimilar words. By instructing adults to use rehearsal, the interaction between tasks and lists disappeared. This finding was in line with the hypothesis that maintenance of verbal information relies on two maintenance mechanisms and adults could choose adaptively between the two, depending on the characteristics of the task or according to instructions.

Experiment 3

In [Experiment 3](#), we instructed participants to use refreshing for maintenance. Assuming that, as in [Experiment 2](#), people would be compliant with instructions regardless of task demand, we predicted that the phonological similarity effect would disappear for both the CRT and the SRT conditions.

Method

Participants

Twenty-four undergraduate students at the University of Bristol (14 women, 10 men) received money for participating. They were all English native speakers, between 18 and 24 years of age ($M = 20.46$, $SD = 1.89$). None of them had participated in the previous experiments.

Materials and procedure

The materials and procedure were the same those as for [Experiment 2](#), except that participants were instructed to use attentional refreshing to maintain the words. We used the same type of instructions as Raye et al. (2007), asking participants to “think of” the words. In the debriefing phase, we questioned participants in order to verify that they had followed instructions. Five participants reporting that they actually had subvocally rehearsed the words were not included in the sample.

Results and discussion

Four participants were discarded from the analysis because they had fewer than 70% correct responses in the CRT task. The pattern of results on the percentages of correct responses and on the RTs for the RT tasks was similar to that in [Experiment 1](#). The percentage of correct responses was higher in the SRT task (95%) than in the CRT tasks (83%), $t(19) = 9.35$, $p < .001$. The one-way ANOVA performed on percentage of correct responses for the CRT task revealed that the three types of lists did not differ

(HS = 82%, HO = 84%, and LO = 84%), $F(2, 38) = 1.14$, $p = .33$, $\eta^2 = .06$. Similarly, for the RTs, the effect of list was not significant, $p = .17$, but the effect of task was significant. The RTs were longer in the CRT task than in the SRT task (496 vs. 265 ms), $F(1, 19) = 73.23$, $p < .0001$, $\eta^2 = .99$. The interaction between task and list was not significant, $p = .29$.

As in the two previous experiments, a 2 (task) \times 3 (list type) ANOVA was performed on the proportion of words recalled in the correct position. The recall performance was poorer for trials with the CRT task than for trials with the SRT task (72% and 85%, respectively), $F(1, 19) = 25.54$, $p < .001$, $\eta^2 = .87$ (Fig. 1). Contrary to what we observed in Experiment 2, the effect of list type was not significant, and neither the difference between the HS list (76%) and the two other types of lists nor that between the HO list (80%) and the LO lists (79%) was significant, $ps > .14$. The interaction between task and type of lists was also not significant, $p = .23$. The lack of a PSE was replicated in the item scores (HS = 85%, HO = 87%, and LO = 87%) and in the order scores (HS = 88%, HO = 91%, and LO = 90%), all $ps > .12$. Finally, the task \times type of lists interaction was also nonsignificant for both the item and order scores, $ps > .32$.

To conclude, when instructed to use refreshing, most participants abandoned articulatory rehearsal and made their best effort to maintain memory traces through refreshing, even in the CRT condition, where this was difficult. As a consequence, the PSE was much diminished and no longer significant, regardless of the attentional demands of the processing task.

Analysis across experiments

To strengthen our conclusion that the phonological similarity effect in a complex span task can be modulated through the instructed maintenance mechanism, we conducted a statistical comparison of Experiments 2 and 3. A 2 (task) \times 3 (list type) \times 2 (instruction) ANOVA on the proportion of words recalled in the correct position showed a significant interaction between list type and instruction, $F(2, 76) = 7.89$, $p < .001$, $\eta^2 = .08$. This interaction reflects the observation that list type (i.e., phonological similarity) had a substantial effect when people were instructed to use articulatory rehearsal (Experiment 2), which was reduced to non-significance when they were instructed to use refreshing (Experiment 3). This interaction was also significant for the item and order scores, $F(2, 76) = 5.83$, $p < .05$, $\eta^2 = .07$, and $F(2, 76) = 4.68$, $p < .05$, $\eta^2 = .12$, respectively.

The comparison of recall performance between Experiment 1 and the other two experiments brought some further evidence on the switching of strategies that we suggested in

Experiment 1. We concluded from the pattern of results observed in Experiment 1 that participants used refreshing in the SRT condition and that they backed up to subvocal rehearsal in the CRT condition, due to a reduction of attentional resources. As a consequence, for the SRT condition, the list type effect should interact with the instructions when Experiments 1 and 2 are compared, because, in the latter experiment, participants used rehearsal. An ANOVA restricted to the SRT condition with instruction (none vs rehearsal) as a between-subjects factor and type of list as a within-subjects factor was performed on the correct-in-position recall, item, and order scores. The instruction \times list interaction was significant for the three scores, $F(2, 70) = 6.86$, $p < .01$, $\eta^2 = .33$, for correct in position, $F(2, 70) = 5.97$, $p < .01$, $\eta^2 = .36$, for item memory, and $F(2, 70) = 4.22$, $p < .05$, $\eta^2 = .35$, for order memory, respectively.

Following the same rationale, in the CRT condition, type of list should interact with instructions when Experiment 1 (no instruction) and Experiment 3 (refreshing) are compared, because we assume that, without instructions, participants opted primarily for rehearsal in the CRT condition. A similar ANOVA was performed on the three scores, showing a significant instruction \times list interaction, $F(2, 70) = 3.36$, $p < .05$, $\eta^2 = .32$, for the correct-in-position score, and $F(2, 70) = 3.35$, $p < .05$, $\eta^2 = .23$, for the item score. This interaction was not significant for the order score, $p = .23$.

General discussion

Previous work has shown that at least two mechanisms contribute to the maintenance of verbal information in the short term: articulatory rehearsal, as described in Baddeley's (1986) model of working memory, and refreshing, as proposed by Johnson (1992) and by Barrouillet and Camos's TBR model (Barrouillet et al., 2004). Our aim in the present study was to evaluate whether the use of these maintenance mechanisms is adaptive, so that it varies according to the relative effectiveness of each mechanism, and whether the choice of maintenance mechanism is under people's intentional control. Because articulatory rehearsal regenerates phonological information, its effectiveness depends on the phonological characteristics of list items. In contrast, refreshing, being a reactivation through attentional focusing, depends on the amount of attention available. We assume that people prefer a maintenance mechanism to the degree that it is effective and, hence, that they should prefer refreshing when attention is available for a relatively large proportion of time, especially in a setting in which many lists are phonologically similar, thus rendering articulatory rehearsal a relatively ineffective form of maintenance. In contrast, they should prefer

articulatory rehearsal when attention is occupied otherwise for most of the time. In addition, we assume a link between the maintenance mechanism used and the kind of representation on which recall relies. Articulatory rehearsal will maintain a phonological representation of the memory list and, hence, lead to a PSE. However, if people engage primarily in refreshing and cease articulatory rehearsal, they are more likely to direct their attention to nonphonological (e.g., semantic, visual) features of the words, and hence, recall will rely primarily on nonphonological representations, so that no phonological similarity effect is to be expected.

In [Experiment 1](#), we manipulated orthogonally the level of phonological similarity of lists of words to be maintained and the attentional demands of the concurrent-processing task in a complex span paradigm. As we expected, the PSE appeared only when the processing component was attention demanding, so that less attention was available for refreshing of the memory traces. Conversely, when the processing requirement was a simple RT task with low attentional demands, the recall of lists of phonologically similar words did not differ from that for lists of dissimilar words. Thus, when enough attention is available and at least some lists are phonologically similar, young adults favor refreshing over rehearsal to maintain words in the complex span paradigm. However, when the amount of attention is reduced, they back up to a less attention-demanding mechanism, articulatory rehearsal. In [Experiment 2](#), we explicitly instructed participants to use articulatory rehearsal. In this case, whatever the amount of attention available, recall was always poorer for lists of phonological similar words, and in addition, an effect of phonological overlap, even in the absence of high similarity, emerged. A different pattern emerged when, in [Experiment 3](#), participants were instructed to use refreshing. Whatever the attentional demands of the concurrent task, the phonological characteristics of the lists to be maintained did not affect recall performance. The comparison of [Experiments 2](#) and [3](#) showed that the PSE depends on the type of maintenance mechanism and the format in which the memory traces are maintained.

The present results converge with those in the work of Anderson (1993) and Siegler (1996) on the choice of strategies. Besides the differences between their views, the commonality of these authors is to conditionalize the choice of strategy on a cost–benefit analysis. In Anderson’s ACT–R model, attention is the resource needed to activate knowledge stored in long-term memory and to implement a process. Thus, to reduce the cognitive cost, a low-demanding strategy should be favored. When it comes to maintenance of verbal information, rehearsal requires a minimal amount of attention, because only the first steps of retrieval of the phonological information for the items

need attention (Baddeley, 2007; Naveh-Benjamin & Jonides, 1984). In contrast, refreshing is highly demanding, because attention needs to be focused on the memory traces throughout the maintenance episodes. Rehearsal and refreshing thus diverge in attentional demands, but they also differ in their efficiency in maintaining materials that could suffer from phonological confusion. Rehearsal allows the maintenance of verbal material only through the reactivation of the phonological characteristics of this material. When lists of words are highly confusable phonologically, refreshing will be more beneficial than rehearsal, because it supports also the maintenance of nonphonological (e.g., semantic) representations of words, which are less confusable. Therefore, with the kind of lists we used, adaptive strategy choice based on a cost–benefit analysis predicts our findings; that is, refreshing should be favored when attention is available, but when attention is restricted, participants should back up to rehearsal.

So far, we have discussed our results in terms of choosing one or the other maintenance mechanism. Could people not use both at the same time? The additive effects of experimental manipulations intended to impair refreshing and to impair rehearsal in previous studies (Camos et al., 2009; Hudjetz & Oberauer, 2007) suggest that they can. The present results leave open the possibility that people can engage in refreshing and rehearsal at the same time, but our data strongly imply that, in the present experiments, participants did not. There is a remarkable consistency across our three experiments. In every condition in which, according to our assumptions, people engaged in refreshing (i.e., the SRT condition in [Experiment 1](#) and both conditions in [Experiment 3](#)), accuracy was approximately 80% in the SRT condition and 70% in the CRT condition. In every instance in which we assumed that people chose to use articulatory rehearsal (i.e., in the CRT condition of [Experiment 1](#) and in all the conditions of [Experiment 2](#)), accuracy dropped below these levels specifically for high-similarity lists. The comparatively good performance in [Experiment 3](#) shows that refreshing was possible to some degree even in the CRT condition. If people had used refreshing to the best of their potential in all the conditions of all the experiments, alone or concurrently with articulatory rehearsal, their performance would never have dropped below the level reached in [Experiment 3](#). The fact that it did—for the HS lists—leads us to conclude that people forfeited refreshing when they opted for rehearsal. Conversely, the fact that the phonological similarity effect disappeared whenever people chose to use refreshing compels the conclusion that in these cases, people abandoned rehearsal, or if they did rehearse, they at least made no use of the rehearsed phonological representations at recall.

A second new finding pertains to the effect of phonological overlap of list items in complex span task. Previous work (Lange & Oberauer, 2005; Oberauer, 2009; Oberauer & Lange, 2008) has demonstrated a detrimental effect of high phonological overlap between memory items and distractors on recall. Here, we investigated for the first time whether a high degree of phonological overlap between all the items on a memory list also would impair memory. The results were mixed: In [Experiment 1](#), overlap had no effect, whereas in [Experiment 2](#), it had an effect, although smaller than the effect of similarity. The two findings are not incompatible, because the instruction in [Experiment 2](#) encouraged the use of phonological representations in all the conditions, whereas in [Experiment 1](#), participants probably used phonological representations only in one condition, so that the power for detecting the relatively small effect of phonological overlap was larger in the second experiment.

A comparison of the present results with those in Oberauer (2009) is instructive. Oberauer (2009) manipulated phonological overlap and phonological similarity between memory words and distractor words to be read aloud in a complex span task, creating three conditions analogous to those in our experiments (i.e., HS, HO, and LO). Whereas high phonological overlap with little similarity (HO) impaired memory, high phonological similarity (HS) did not. In the present experiments, high similarity had a larger detrimental effect than did high overlap with little similarity. The difference between the studies can be explained as follows: Items on the memory list are all treated as candidates for recall; therefore, when list items are similar to other list items, people can easily confuse the list item to be recalled at a given serial position with another list item. Distractors in a complex span task, in contrast, are not necessarily treated as candidates for recall. People know that distractors should not be remembered, so they can exclude them from the candidate set by (1) avoiding encoding them into working memory (Awh & Vogel, 2008), (2) removing them from working memory soon after encoding (Oberauer, 2001), or (3) marking them as not to be recalled by association to a separate context (Delaney & Sahakyan, 2007). To the degree that one of these three processes successfully excludes distractors from the candidate set, people are unlikely to mistakenly recall a distractor instead of an item, even when the distractor is highly similar to the list item. This explains why phonological similarity between memory items impairs memory, whereas phonological similarity between memory items and distractors does not. At the same time, phonological overlap seems to impair memory regardless of whether overlap occurs between memory items or between a memory item and a set of distractors. This is to be expected from the assumed

mechanism generating the effect, feature overwriting. Any representation encoded into working memory, either as a memory item to be recalled later or as a distractor to be processed and then discarded, has the potential to overwrite shared features of memory items. At the same time, overwriting of phonological features matters only insofar as memory depends on phonological representations. Therefore, under conditions in which people choose refreshing over rehearsal, so that nonphonological representations become more important for recall, the effect of phonological overlap is reduced for the same reasons that the effect of phonological similarity is reduced.

To conclude, the present study presents evidence that adults can choose between articulatory rehearsal and attention refreshing to maintain verbal information in working memory. This choice depends on instructions, but also on the attentional demands of concurrent processing. As is suggested by the extended TBRS model, relying on the peripheral level—that is, maintaining phonological codes through articulatory rehearsal—enhances the impact of the phonological characteristics of the material to be maintained. Conversely, the maintenance of nonspecific representations through refreshing diminishes the PSE. Differences between experiments in task characteristics that encourage the use of one or the other maintenance mechanisms, as well as differences between participant samples with different strategic preferences, could explain why previous studies yielded divergent findings on the PSE in complex span tasks.

Acknowledgment We thank Helen Tam and Tim Jones for their contribution in the selection of the materials and Simon Farrell for helping us to organize testing session. Part of this work was done when V.C. was invited fellow at University of Bristol funded by the Leverhulme Trust. G.M. is funded by a grant from Ministère de la Recherche (Bourse MENRT). K.O. was supported by a grant from the Economic and Social Research Council (ESRC), Grant RES-062-23-1199. K.O. is now at the University of Zurich.

References

- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale: Erlbaum.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.
- Awh, E., & Vogel, E. K. (2008). The bouncer in the brain. *Nature Neuroscience*, *11*, 5–6.
- Baddeley, A. D. (1966a). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly Journal of Experimental Psychology*, *18*, 302–309.
- Baddeley, A. D. (1966b). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity.

- The Quarterly Journal of Experimental Psychology*, 18, 362–365.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (2007). *Working memory, thought, and action*. New York: 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, 83–100.
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 570–585.
- Barrouillet, P., & Camos, V. (2001). Developmental increase in working memory span: Resource sharing or temporal decay? *Journal of Memory and Language*, 45, 1–20.
- Barrouillet, P., & Camos, V. (2007). The time-based resource sharing model of working memory. In N. Osaka, R. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59–80). Oxford: Oxford University Press.
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, 61, 457–469.
- Campoy, G., & Baddeley, A. (2008). Phonological and semantic strategies in immediate serial recall. *Memory*, 16, 329–340.
- Cohen, J. D., Mac Whinney, B., Flat, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, & Computers*, 25, 257–271.
- Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, 55, 75–94.
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology*, 55, 429–432.
- Copeland, D. E., & Radvansky, G. A. (2001). Phonological similarity in working memory. *Memory & Cognition*, 29, 774–776.
- Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.
- Cowan, N. (2005). *Working memory capacity*. Hove: Psychology Press.
- Delaney, P. F., & Sahakyan, L. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory & Cognition*, 35, 1074–1082.
- Fallon, A. B., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When CAT helps RAT, but not MAN. *International Journal of Psychology*, 34, 301–307.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9, 59–79.
- Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 838–849.
- Fournet, N., Juphard, A., Monnier, C., & Roulin, J.-L. (2003). Phonological similarity in free and serial recall: The effect of increasing retention intervals. *International Journal of Psychology*, 38, 384–389.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove: Psychology Press.
- Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review*, 10, 435–444.
- Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory & Cognition*, 35, 1675–1684.
- Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, 4, 268–280.
- Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective & Behavioral Neuroscience*, 5, 339–361.
- Lange, E. B., & Oberauer, K. (2005). Overwriting of phonemic features in serial recall. *Memory*, 13, 333–339.
- Lobley, K. J., Baddeley, A. D., & Gathercole, S. E. (2005). Phonological similarity effects in verbal complex span. *The Quarterly Journal of Experimental Psychology*, 58A, 1462–1478.
- Macken, W. J., & Jones, D. M. (2003). Reification of phonological storage. *The Quarterly Journal of Experimental Psychology*, 56A, 1279–1288.
- Mora, G., Barrouillet, P., & Camos, V. (2009). A developmental investigation of maintenance mechanisms in working memory through the phonological similarity effect. Communication at the *European Society of Cognitive Psychology Meeting*, Krakow.
- Murray, D. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology: General*, 78, 679–684.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, 18, 251–269.
- Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 369–385.
- Oberauer, K. (2001). Removing irrelevant information from working memory. A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 948–957.
- Oberauer, K. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition*, 37, 346–357.
- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55, 601–626.
- Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: Distinguishing similarity-based confusion, feature overwriting, and feature migration. *Journal of Memory and Language*, 58, 730–745.
- Raye, C. L., Johnson, M., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex*, 43, 135–145.
- Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a single thought: Dorsolateral PFC activity associated with refreshing just-activated information. *Neuroimage*, 15, 447–453.
- Romero, L., Walsh, V., & Papagno, C. (2006). The neural correlates of phonological short-term memory: A repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 18, 1147–1155.
- Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 150–164.
- Salamé, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society*, 24, 263–265.

- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Sperling, G. (1967). Successive approximations to a model for short term memory. *Acta Psychologica*, *27*, 285–292.
- Tehan, G., Hendry, L., & Kocinski, D. (2001). Word length and phonological similarity effects in simple, complex and delayed serial recall tasks: Implications for working memory. *Memory*, *9*, 333–348.
- Tehan, G., & Hymphreys, M. S. (1998). Creating proactive interference in immediate recall: Building a DOG from a DART, a MOP, and a FIG. *Memory & Cognition*, *26*, 477–489.
- Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Verbal and visuo-spatial working memory: A case for domain-general time-based resource sharing. *Psychological Science*, *21*, 384–390.
- Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, *72*, 89–104.