The Role of Semantic Representations in Verbal Working Memory

Vanessa M. Loaiza

Valérie Camos

Word count: 15,359 (excluding references)

Author Note

Vanessa M. Loaiza is now at Department of Psychology, University of Essex.

Valérie Camos, Department of Psychology, Fribourg Center for Cognition, University of Fribourg.

Correspondence concerning this article should be addressed to Vanessa M. Loaiza, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: v.loaiza@essex.ac.uk

This research was funded by the Swiss National Science Foundation grant awarded to Valérie Camos, 100014_137860. We acknowledge Loïc Apreda, Julie Bessire, Claudie Gaillard, Kevin Moser, Benoît Perriard, and Vania Sandoz for their assistance with data collection.

Abstract

Word count: 248

Two main mechanisms, articulatory rehearsal and attentional refreshing, are argued to be involved in the maintenance of verbal information in working memory (WM). While converging research has suggested that rehearsal promotes the phonological representations of memoranda in working memory, little is known about the representations that refreshing may promote. Not only would examining this question address this gap in the literature, but the investigation has profound implications for different theoretical proposals of how refreshing functions and on the relationships between WM and long-term memory (LTM). Accordingly, we tested predictions from five models regarding how refreshing may moderate the semantic representation of memoranda in verbal WM. This series of four experiments presented a cue word that was either semantically or phonologically related to a target during the recall phase of a complex span task. Experiment 1 established the benefit of semantic over phonological retrieval cues, and Experiment 2 established that this semantic benefit was specific to a refreshing- rather than a rehearsal-based maintenance strategy. Finally, we showed that this semantic benefit did not vary with the cognitive load of the concurrent task (Experiments 3 and 4) or the intention to learn the memoranda (Experiment 4). These results indicate that cue-based retrieval from episodic LTM may strongly contribute to semantic processing effects in WM recall, but this influence of episodic LTM is independent of the function of refreshing to reactivate memory traces. Accordingly, these results have strong implications for the functioning of refreshing and the links between WM and LTM.

Keywords: working memory, attentional refreshing, semantic cues, long-term memory

The Role of Semantic Representations in Verbal Working Memory

Much research has been devoted to understanding the underlying mechanisms that support the temporary and limited maintenance of presently active verbal information, or storage in working memory (WM). Early WM models focused on the maintenance of verbal information via articulatory rehearsal (Baddeley, 1986), but more recent models have investigated an additional mechanism, attentional refreshing, that is purported to be qualitatively distinct in its functioning (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Camos, Lagner, & Barrouillet, 2009; Camos, Mora, & Barrouillet, 2013; Camos, Mora, & Oberauer, 2011). Whereas articulatory rehearsal (or simply, rehearsal) is thought to operate by subvocal, covert, and phonologically-based repetition of the memoranda, attentional refreshing (or simply, refreshing) is considered a domain-general, attention-based mechanism that operates by briefly thinking back to recently active memoranda (see Camos, 2015, 2017 for review). This recent research has demonstrated that the two maintenance mechanisms are independent in their effects on verbal WM recall (Camos et al., 2009), particularly considering the characteristics of the memoranda (Camos et al., 2013, 2011; Mora & Camos, 2013). More specifically, much of this research has focused on how rehearsal, and not refreshing, may promote the phonological characteristics of verbal memoranda. However, how refreshing operates and which characteristics it may promote is much less understood. The following experiments addressed this gap in the literature using a novel paradigm in order to comprehensively examine several predictions regarding the impact of refreshing on semantic representations in verbal WM. Accordingly, this study sheds light on how refreshing functions and more generally on the relationship between WM and long-term memory (LTM).

Verbal maintenance in working memory

Early models concerning WM and short-term retention of verbal information focused on the important role of articulatory rehearsal to briefly maintain and keep information active (Baddeley, 1986). Much of the evidence for rehearsal as the principal mechanism for verbal WM relied on manipulations that varied the phonological status of the memoranda (Baddeley, 1966; Conrad, 1964; Conrad & Hull, 1964) or their relative ability to be articulated (Baddeley, Thomson, & Buchanan, 1975; Levy, 1971). These studies were the first to demonstrate that rehearsal is a domain-specific mechanism that emphasizes the phonological characteristics of the memoranda. For example, the phonological similarity effect (Baddeley, 1966; Conrad, 1963) or the word length effect (Baddeley et al., 1975) are well-replicated effects that phonologically similar memoranda (e.g., *mad, man, cap, cat*) or longer words (e.g., *association, opportunity, representative*) are less likely to be recalled than dissimilar memoranda (e.g., *cow, day, bar, few*) or shorter words (e.g., *sum, hate, harm*) during tests of immediate serial recall, respectively. Thus, when memoranda are more phonologically confusable or take longer to articulate, rehearsal is less efficient to maintain the memoranda in WM because it relies on their phonological characteristics. This was considered especially evident with findings that these effects disappeared when rehearsal of visually presented memoranda was blocked through concurrent overt articulation (i.e., articulatory suppression; Baddeley et al., 1975; Levy, 1971).

Importantly, maintenance of verbal information in Baddeley's original model relies on the domain-specific mechanism of rehearsal, whereas attention has no role in maintenance (see also Logie, 2011; although Baddeley, 2012, recently acknowledged a possible role of attentional refreshing in the episodic buffer). However, some have further argued for an additional, domain-general maintenance mechanism in WM called attentional refreshing. There have been several different proposals of how refreshing operates since its first conception, and while there are some similarities across them, thus far there has not been a consensus in the literature. Johnson (1992) first established refreshing as a process that serves to prolong the activation of information by reflectively thinking back to its just previously activated representation. Similarly, in Cowan's (1999) embedded processes model, because WM is the activated part of LTM, information must be recirculated through the focus of attention in order to keep them active. More recently, Vergauwe and Cowan (2014, 2015) have suggested that refreshing may function as a scanning or search of the central component of WM. In a complex span task, this was evidenced by the absence of detrimental effect of the concurrent task when it was a memory search task (decide whether the distracting letter was represented in the memory set so far), compared to the detrimental effect of other tasks, such as a location (decide whether the letter was up or down on the screen) or alphabet judgment task (decide whether each letter came before or after the letter O in the alphabet).

Other models have also emphasized the importance of refreshing in WM. The time-based resource sharing (TBRS) model of WM has also espoused refreshing via attentional focusing (Barrouillet et al., 2004; Barrouillet, Portrat, & Camos, 2011). In this model, refreshing is a purposeful mechanism of keeping memoranda in WM active by focusing attention to their representations, especially after a period of distraction where their activation may have decayed. However, because attention is a limited resource that can only be allocated to one activity at a time, it has to switch between maintenance and processing activities during WM tasks. As a consequence, the TBRS model predicts that maintenance by refreshing is specifically limited by the cognitive load (or attentional demand) of the processing activities. For example, during a complex span task, participants are instructed to briefly maintain a series of memoranda (e.g., words) that are each followed by a series of processing episodes to respond to (e.g., several successively presented black squares that appear in the upper or lower part of the screen). In such a task, under equivalent timing conditions, processing decisions that are more attention-demanding (e.g., responding as to whether a square is up or down on the screen, i.e., a choice-reaction time task) would distract attention for a longer period of time, hence reducing the ability to refresh memoranda relative to less attention-demanding decisions (e.g., pressing a key for each square's appearance, i.e., a serial reaction time task), and leading to poorer recall performance (Barrouillet et al., 2007; Barrouillet et al., 2011; Camos et al., 2009). Thus, manipulating the cognitive load of a processing task has proven to be an important factor that determines WM capacity, possibly due to the specific manipulation of refreshing.

McCabe's (2008) covert retrieval model similarly draws upon the importance of refreshing in WM. Following Unsworth and Engle's (2007) primary-secondary memory framework, the covert retrieval model suggests that about four chunks of information can be maintained in primary memory, but new incoming memoranda or a distracting processing task displace them into secondary memory (i.e., LTM). Consequently, the memoranda must be retrieved from LTM in order to keep them accessible for WM recall. Rose and Craik (2012) have similarly asserted that the extent to which LTM factors (e.g., levels-of-processing effects) contribute to WM performance depends on the amount of disruption to active maintenance processes. More recent research supporting this view has suggested that refreshing is this act of retrieval from LTM, such that attention must be used to refresh the memoranda from a less active state (Loaiza & McCabe, 2012, 2013). In these studies, Loaiza and McCabe (2012, 2013) varied the placement of the interleaving processing task amongst the memoranda to vary the opportunities to purportedly refresh the memoranda. For example, in an operation span task that alternates presentation of memoranda with arithmetic problems, each memorandum must be refreshed after each arithmetic problem to keep it active, and thus each memorandum has successively fewer opportunities to be refreshed as the trial progresses. Congruent with the covert retrieval model, performance during a delayed test, which assessed retrieval from episodic LTM, increases with the number of refreshing opportunities (Loaiza & McCabe, 2012, 2013; Loaiza, Rhodes, & Anglin, 2015).

Taken together, these studies suggest that attentional refreshing is an important mechanism to sustain the activation of memoranda in WM. Furthermore, its method of maintaining information via attention rather than more peripheral mechanisms like rehearsal suggests that its functioning is qualitatively distinct from rehearsal, as we will see in the next section. It should be noted that all the previously presented models have espoused that verbal information is temporarily maintained in WM, which explains the need for some maintenance mechanisms. However, alternative views put less emphasis on active maintenance.

The unitary view of memory (e.g., Crowder, 1982; Melton, 1963; Nairne, 1990; Nairne, 2002) argues that, much like episodic LTM, retrieval over the short term is cue-based and does not rely on any postulated maintenance mechanisms. This would mean that there is no need to posit a refreshing mechanism much less a WM system to account for memory over the short term. As such, any factors presumed to affect episodic LTM should similarly affect WM, regardless of any use of refreshing or rehearsal to actively maintain the memoranda. For example, using semantic cues that benefit memory performance by drawing upon pre-existing semantic associations in LTM (Howard & Kahana, 2002) should similarly improve retrieval from WM, irrespective of the intention or strategies to actively maintain the memoranda. Besides this critical difference regarding active maintenance, such a view also strongly differs with the previously described WM models concerning the relationship between WM and LTM, as it sees no need in dissociating the two. Conversely, WM models hypothesize a WM, but differ on how distinct WM is from LTM: from totally separable (e.g., the multicomponent and TBRS models) to WM as the activated part of LTM (e.g., the embedded processes model). Thus, not only are there different theoretical conceptions of how verbal maintenance is achieved, but the question also brings much to bear regarding the distinction or relationship between memory systems like WM and LTM. We will revisit this topic again after first outlining the relevant literature regarding verbal maintenance and semantic representations in WM.

Two qualitatively distinct mechanisms

Recent research from the TBRS model has justified this prediction that attentional refreshing and articulatory rehearsal are distinct maintenance mechanisms that operate jointly but independently to support WM recall (Camos et al., 2013, 2011; Hudjetz & Oberauer, 2007; Mora & Camos, 2013). Camos and colleagues (2009) demonstrated that manipulating the opportunity for rehearsal or refreshing had independent and additive effects on WM recall. Refreshing and rehearsal could be manipulated by increasing the cognitive load (i.e., attentional demand) of the task or by requiring concurrent articulation (i.e., articulatory suppression) during the task, respectively. For example, WM recall was reduced when manipulating refreshing through increased cognitive load while also controlling for rehearsal by requiring concurrent articulation during the WM task (Camos et al., 2009). This demonstrated the existence of refreshing as a maintenance mechanism that could operate even when rehearsal was blocked by articulatory suppression. Likewise, holding refreshing constant but manipulating rehearsal also dramatically reduced WM recall (Camos et al., 2009). Camos and colleagues (2009) also showed that orthogonally manipulating refreshing and rehearsal yielded additive effects of both mechanisms, and the factors themselves did not interact. The results demonstrated that both refreshing and rehearsal contribute to WM recall, but do so independently. These behavioral results are corroborated by neuroscientific studies showing that refreshing and rehearsal are subserved by different neural correlates, such that the dorsolateral prefrontal cortex is uniquely active when participants must refresh a word relative to re-reading (Raye, Johnson, Mitchell, Reeder, & Greene, 2002) or rehearsing a word (Raye, Johnson, Mitchell, Greene, & Johnson, 2007). Furthermore, the distinction is also evident in episodic LTM (Camos & Portrat, 2015; Loaiza & McCabe, 2013; Rose, Buchsbaum, & Craik, 2014) and in the aging literature (Loaiza & McCabe, 2013; Raye, Mitchell, Reeder, Greene, & Johnson, 2008). Thus, converging evidence from multiple domains of research has supported the distinction between rehearsal and refreshing in WM maintenance.

Given this distinction between rehearsal and refreshing, Camos and colleagues (2011; 2013; Mora & Camos, 2013) have further investigated whether the qualitative nature of WM recall also differs between rehearsal and refreshing. That is, the prediction that rehearsal is especially important for phonological representations in verbal WM (Baddeley, 1966) implies that its efficiency for WM maintenance is more susceptible to the phonological characteristics of the memoranda. Conversely, refreshing is considered as a domain-general and attention-based mechanism, and thus the phonological status of the memoranda should not moderate its effect on WM recall. A series of studies demonstrated just this. For example, Camos and colleagues (2013) showed that the phonological similarity effect was evident when it was possible to use rehearsal during a complex span task, but not when rehearsal was suppressed. Conversely, the effect was evident when cognitive load was high (Camos et al., 2013; 2011), supporting the notion that participants are less able to use refreshing under high attentional demands and therefore must flexibly switch to rehearsal to maintain the memoranda. Likewise, the effect disappeared when participants were specifically instructed to use refreshing to maintain memoranda, whereas it was present when participants were instructed to use rehearsal, even under high cognitive load conditions (Camos et al., 2011). Mora and Camos (2013) showed similar results with the word length effect: articulatory suppression eliminated the word length effect, whereas the effect was still present even when the cognitive load of the task increased. These findings demonstrate the flexibility of use between rehearsal and refreshing and support their distinguishability by indicating that the nature of the representations in WM are qualitatively distinct as a function of using either maintenance mechanism. That is, using rehearsal appears to both rely upon and emphasize the phonological characteristics of the memoranda in WM.

Semantic representations in working memory

This begs the question: if articulatory rehearsal reinforces the phonological characteristics of the memoranda, then what is the nature of the characteristics that refreshing promotes? Given their separable effects on WM recall, it is probable that refreshing and rehearsal also emphasize qualitatively different characteristics of memoranda in WM. However, to date, the research concerning this question is sparse. Camos et al. (2011) argued that relying on refreshing to maintain verbal memoranda prompts attention toward non-phonological features of the memoranda, such as semantic characteristics. Accordingly, rehearsal and refreshing may not just be distinguishable in terms of their effects on WM recall, but also to the degree that they differentially emphasize phonological or semantic characteristics of the memoranda, respectively.

In the related domain of short-term memory (STM), patient as well as neuroimaging studies bring convergent evidence in favor of two distinct neural networks underlying verbal STM (for a review, see Martin, 2005). For example, Hanten and Martin (2000) distinguished one network subserving the retention of phonological information, which involves the superior temporal lobe and the supramarginal gyrus, from another that maintains semantic information by recruiting the inferior and middle temporal lobe and the inferior frontal lobe. Accordingly, Hamilton, Martin, and Burton (2009) reported cases of patients with damage in the left inferior and middle frontal gyri who exhibited deficits in semantic STM, whereas patients with lesions in the inferior parietal areas exhibited phonological STM deficits. Further studies with normal participants have also demonstrated similar double dissociations (Nishiyama, 2013, 2014). For example, using a dual-task interference paradigm, Nishiyama

(2014) showed that disrupting maintenance of words using articulatory suppression or tapping differently affected performance on a homophone (phonological) or synonym (semantic) short-term recognition task. Whereas reducing the participants' ability to engage in articulatory rehearsal reduced performance on the homophone task relative to tapping, the opposite was true for the synonym task, such that the more attention-demanding tapping reduced performance on the synonym task relative to articulatory suppression. These studies highlight that verbal maintenance in STM relies on two distinct networks promoting either phonological or semantic representations. Shivde and Anderson (2011) made a similar proposal for WM, such that maintenance of semantic representations is independent from phonological representations in WM. Their series of studies used a concurrent probe method, such that participants were presented with a to-be-maintained word and instructed to attend to its meaning or phonology for a later probe decision. Critically, a lexical decision task filled the interval between the presentation of the to-be-maintained word and the probe, wherein either a semantically related or unrelated word was presented amongst other words and non-words. Their results showed that participants responded to semantically related words more slowly when they were instructed to attend to the meaning versus the phonology of a target word, whereas the opposite pattern was shown for phonologically related words (Shivde & Anderson, 2011). Together these studies suggest that semantic and phonological representations are distinguishable in STM and WM.

Relatedly, a growing literature has examined the role of traditionally investigated semantic effects in the LTM literature in WM paradigms. Such effects include the lexicality effect (i.e., that words are better recalled than non-words; Loaiza, Duperreault, Rhodes, & McCabe, 2015), the frequency effect (i.e., that highly frequent words are better recalled than low frequency words, Engle, Cantor, & Carullo, 1992), and the concreteness effect (i.e., that concrete, imageable words are better recalled than abstract, low-imageability words;

Campoy, Castella, Provencio, Hitch, & Baddeley, 2015). These studies have demonstrated that these LTM effects are likewise evident in WM tasks. Perhaps the LTM effect that has received the most recent attention in WM is the levels-of-processing effect (Craik & Tulving, 1975). These studies have suggested that a deeper, more semantically meaningful method of studying memoranda yields greater WM recall than shallowly studying memoranda during complex span tasks that distract attention from maintenance of memoranda (Loaiza, McCabe, Youngblood, Rose, & Myerson, 2011; Rose et al., 2014; Rose & Craik, 2012; Rose, Craik, & Buchsbaum, 2015; Rose, Myerson, Roediger, & Hale, 2010), just as in the traditional levels-of-processing effect regularly shown in episodic LTM (Craik & Tulving, 1975). This strengthens the idea that semantic representations, in addition to phonological representations, are evident in verbal WM.

Congruent with this idea, other paradigms have investigated the role of semantic representations in WM by considering semantic interference. For example, Atkins and Reuter-Lorenz (2008) successively presented four semantically related memoranda that were either followed by a distracting arithmetic problem or a simple reaction time task for the same fixed duration. Depending on the experiment, retrieval from WM was tested either by recall or item recognition, wherein a probe word was presented that was either in the memory set (i.e., a positive probe), a never-presented and unrelated word (i.e., a negative probe), or a semantically related probe (i.e., a lure probe). Just as in the episodic LTM literature, Atkins and Reuter-Lorenz demonstrated false memory in WM, such that semantically-related intrusion probes were more likely to be falsely recognized and were slower to reject during retrieval from WM (see also Flegal, Atkins, & Reuter-Lorenz, 2010; Flegal & Reuter-Lorenz, 2014). Furthermore, false recall of semantic lures occurred more often than phonological intrusions and other unrelated intrusions. Importantly, these semantic intrusion effects were greater when a more attentionally demanding distracter filled the retention interval between presentation of the memoranda and recall compared to the simple reaction time task (Atkins & Reuter-Lorenz, 2008). As explained previously, the TBRS model has conceptualized such variations in the cognitive load of a concurrent task as specifically affecting the efficiency of refreshing in WM. Although Atkins and Reuter-Lorenz did not interpret this finding in detail or appeal to the TBRS model, the results may suggest that the processing of semantic representations of the memoranda is sensitive to manipulations of cognitive load. That is, increasing the cognitive load of the concurrent task, and thus presumably affecting the efficiency of refreshing, likewise increased the rate of semantic interference whereas other intrusion errors were relatively stable. Higgins and Johnson (2013) further demonstrated that semantic interference specifically affected the efficiency of refreshing information relative to unrelated distractors. This was shown through implicit interference of briefly presented masked distracters that were related or unrelated to target words that were either repeated or refreshed immediately after their presentation. As the authors expected, participants refreshed the target words significantly more slowly during the related than unrelated trials, though there was no difference for the repeated words (Higgins & Johnson, 2013). This shows that items in the focus of attention or currently refreshed could be sensitive to semantic interference. Conversely, presenting semantically-related distractors after each memory item in a complex span task improved recall performance, an effect that disappeared under rehearsal, but that was not affected by variations in cognitive load (Oberauer, 2009).

Although the paradigms differ between studies, the results collectively suggest that semantic processing is evident in WM, but it remains to be demonstrated whether refreshing may be sensitive to semantic representations. Some studies have shown that semantic effects may vary with the opportunity to engage in refreshing, whether by varying the cognitive load (Atkins & Reuter-Lorenz, 2008) or directing participants to refresh (Higgins & Johnson, 2013), whereas others do not (Oberauer, 2009). By contrast, rehearsal does not appear to be sensitive to the semantic characteristics of the memoranda. However, thus far these studies have not explicitly compared the nature of the characteristics that are differently emphasized as a function of refreshing. The current study contributes to the literature by directly testing the possible link between refreshing and semantic processing in WM. Moreover, the current study also distinguishes between several conceptions of refreshing in how that link may manifest when instructing rehearsal or refreshing as a strategy to maintain the memoranda, when manipulating the cognitive load of the secondary task, and when encoding is intentional versus incidental. Finally, this study would also yield strong implications regarding the broader issue of the distinction between WM and LTM.

The present study

The goal of the present study was to examine the nature of maintained representations in verbal WM. In particular, we investigated whether semantic versus phonological characteristics of the memoranda are important to WM recall, especially as a function of refreshing, in order to test five different models (Table 1). According to Baddeley's original multicomponent model, there is no attention-based maintenance, and thus verbal maintenance is primarily phonological in nature due to the domain-specificity of the phonological loop (Baddeley, 1986; see also Logie, 2011). By contrast, the TBRS and covert retrieval models both espouse a role of attentional refreshing in WM and consequently both posit a role of semantic processing in WM, but for different reasons. For the TBRS model, memory traces are constructed in WM using information from LTM, and refreshing operates by reconstructing decaying memory traces in an episodic buffer (Barrouillet & Camos, 2015). Thus, semantic processing should be evident in WM recall, especially when refreshing is instructed or available (e.g., during low cognitive load conditions) for use. The covert retrieval model (McCabe, 2008) would likewise predict an effect of semantic processing, but rather because refreshing operates as a covert retrieval from LTM to bring back memory traces in primary memory. Accordingly, reducing the opportunity to refresh information in WM (i.e., in high cognitive load condition) increases the importance of deep, semantic processing due to an increased reliance on resources from episodic LTM to support performance (Rose et al., 2014; Rose & Craik, 2012; Rose et al., 2015). Still another view has been recently put forward within the embedded processes model in which refreshing may operate as a rapid search of the active content in WM (Vergauwe & Cowan, 2014, 2015). In this regard, semantic processing would have an overall effect on WM due to its embeddedness within LTM, but this effect would not vary with the availability of refreshing. Finally, a unitary view of memory (e.g., Nairne, 2002) would posit that semantic processing has an overall effect on memory, regardless of refreshing, because there is no functional distinction between WM and episodic LTM. The following series of experiments aimed to elucidate the nature of semantic processing in verbal WM in order to address a deep asymmetry of the literature's focus on rehearsal and phonological representations. However, even more importantly, the study also helps to adjudicate between different models of WM and their conceptions of refreshing.

In order to address these predictions, the following series of experiments introduced a novel paradigm to explore whether semantic retrieval cues are particularly helpful when participants forget memoranda during WM recall. Specifically, we used a complex span task that presented five target memoranda (e.g., *bread*) interleaved by a distracting location judgment task (e.g., deciding whether a square was presented up or down on the screen). The advantage of this task is that it is already known to impede upon attention-based maintenance (e.g., Barrouillet et al., 2004, 2007). The other advantage is that it also not a verbal task, and thus we can be sure the effects are due to constraints on attention rather than representation-based interference. During recall, participants were presented with either phonological (e.g., *thread*) or semantic (e.g., *sandwich*) retrieval cues for each forgotten target memorandum (Experiments 1-3) or all of the memoranda (Experiment 4).

To foreshadow, Experiment 1 first established that there is a benefit of receiving semantic retrieval cues relative to rhyme cues, henceforth referred to as the semantic retrieval cue benefit. Experiment 2 manipulated the instructions given to participants to maintain the memoranda via rehearsal or refreshing to examine whether the semantic retrieval cue benefit is specific to refreshing. Experiment 3 examined whether manipulating the cognitive load of the distracting processing component, and thereby the efficacy of refreshing, moderated the semantic retrieval cue effect. Finally, Experiment 4 manipulated the cognitive load and whether the memory test was expected (i.e., incidental versus intentional encoding) in order to address whether active maintenance in WM moderates the semantic retrieval cue benefit.

There are specific predictions for each experiment given the five different tested models (Table 1). In general, the original multicomponent model assumes that verbal maintenance is primarily phonological, and thus a phonological cue benefit would be evident across experiments, except when active maintenance is not instructed (i.e., during incidental encoding) and retrieval therefore requires episodic LTM. The TBRS model predicts that a semantic retrieval cue benefit should be most evident when refreshing is instructed or available (i.e., during low cognitively demanding activities) during active maintenance of memoranda in WM (i.e., intentional encoding). Conversely, the covert retrieval model would predict that a semantic retrieval cue benefit should be most evident when refreshing is instructed or required to retrieve displaced memoranda from episodic LTM (i.e., during high cognitively demanding activities) during active maintenance of memoranda in WM (i.e., intentional encoding). The embedded processes model would posit that a semantic retrieval cue benefit should be evident regardless of the availability of refreshing during active maintenance of memoranda in WM. Finally, if episodic LTM strongly contributes to WM recall, such that they may not even be functionally distinct systems, then a semantic retrieval cue benefit should be evident regardless of factors intended to manipulate active maintenance WM, thus supporting a unitary view of memory.

Experiment 1

Experiment 1 established whether there is an overall benefit of semantic over phonological retrieval cues when recall from WM fails. Baddeley's original multicomponent model would oppositely predict that phonological cues should be more helpful to WM recall than semantic cues given that verbal maintenance is achieved by rehearsal and thereby more phonological in nature. Conversely, the TBRS, covert retrieval, and embedded processes models would predict a benefit of semantic over phonological cues due to the importance of refreshing to promote semantic processing in WM. A unitary model of memory would also posit that semantic retrieval cues should be beneficial to memory overall. Participants were presented with two successive blocks of complex span trials, during which semantic or phonological cues were presented (depending on the block) if they forgot any memoranda and asked for help.

Method

Participants. Thirty-three participants were recruited from the University of Fribourg in exchange for partial course credit or cinema ticket coupons. The data from eight participants were excluded due to the fact that these participants did not use the retrieval cues at least 12% of the time in either block (see Design and Scoring). This left 25 participants for the analyses (22 female, M_{age} = 22.40, *SD* = 6.37). All participants were native French speakers. All participants in each of the experiments of this study provided informed consent before beginning the experiment and were debriefed at the end of the experiment. The ethics committees of the University of Fribourg (Experiments 1 – 3) and the University of Essex (Experiment 4) approved the ethics applications for the study.

Materials and Procedure. In Experiments 1 to 3, 55 highly frequent, concrete words served as the memoranda and were selected from the French *Lexique* database (New, Pallier, Ferrand, & Matos, 2001). Five of the memoranda were used for a practice trial. Each memorandum was associated to a phonological and semantic cue word. Phonological cue words were selected such that central and final phonemes were the same for the target words and their respective phonological cues (e.g., *moto* and *photo*). Semantic cue words were selected from norms developed in French by Ferrand and Alario (1998). They were selected such that the forward associative strength between the semantic cue and target word was between 15-49% (M = 29.58, SD = 8.83). This range was selected so as to ensure a semantic association between the semantic cues and targets without encouraging guessing. Subsequent analyses suggested that this was the case, as the correlations between semantic strength and cue accuracy were positive but low and non-significant across Experiments 1-3 (rs < .23, ps > .09). Errors due to incorrect guesses in response to the cues in the semantic condition were also low overall (M = 28%, SD = 33% of errors across Experiments 1-3; i.e., most errors were omission errors). Furthermore, the semantic strength of the incorrect guesses to the cues (as indexed by the Ferrand and Alario norms) was also very low (M = 3.03%, SD = 4.88% across Experiments 1-3). Overall, these analyses suggest that the chosen semantic strength was satisfactory to discourage guessing merely on semantic strength. Target words were pre-arranged into trials in order to ensure that there was no overlap between the words in terms of phonological or semantic relatedness (i.e., that the cues were specific to that word in the trial). The memoranda were counterbalanced for cue condition across participants.

The experiment session began with a practice phase of the location judgment task. Participants were presented with 20 black squares presented successively and randomly in the upper or lower quadrant of the screen. To discourage rehearsal, participants had to respond aloud and press one of two designated keys on the keyboard as to whether each square was up or down on the screen. Response times (RTs) from their key presses were collected. Each square was presented for 700 ms with a 300 ms interstimulus interval (ISI). Participants were required to reach an 85% criterion to pass to the next phase of the experiment; those who did not repeated the practice phase until reaching the criterion.

Participants then completed two blocks corresponding to the phonological or semantic cues, with five trials per block. The block order was counterbalanced and trial presentation was random, and each block began with instructions and an example trial. Each trial began with a fixation cross at the center of the screen for 750 ms, immediately followed by a target word to remember presented in red font for 1 s. After a 500 ms ISI, participants again saw the squares appear in the upper or lower part of the screen. The position of the square was random. As in the practice task, participants said “up” or “down” while pressing the corresponding key on the keyboard. There were four squares successively presented for 700 ms with a 300 ms ISI separating the presentation of each square. After the location judgment decisions, another target word was presented. This sequence repeated five times in the trial for five total words to remember by the end of the trial. The trial ended with a screen that said “Rappel!” (*recall* in English) with numbers 1-5 in a column representing the target words’ serial position in the trial. Participants were instructed to type the target words next to the numbers representing the original order in which they were presented. Note that the phonological and semantic blocks were thus identical to this point. The participants were instructed that if they had forgotten a word after trying to remember it, they were allowed to ask for help from the experimenter. When the participants asked for help, depending on the block, a phonological or semantic cue word appeared next to the number representing that target word’s serial position. Participants then could try to recall the word based on the presentation of the cue. The participants were instructed to use the help when they really needed it. The duration of the experiment was approximately 30 min.

Design and Scoring. There were three dependent variables: recall, cue use, and cued recall accuracy (i.e., the likelihood of correct recall if the participants asked for a cue). The independent variable was the type of cue (phonological or semantic). Each dependent variable was assessed using three separate paired-samples *t*-tests comparing the phonological and semantic blocks.

Due to the nature of the cues being provided at the request of the participant, in all three experiments, there were some participants who never or very rarely used the retrieval cues. This occurred either because of perfect recall, transposition or commission errors that resulted in inaccurate serial recall but also no request for a cue, or more rarely, they did not recall a word at all or ask for a cue (i.e., they skipped to the next word to recall). The latter was discouraged by the experimenter, but because participants had control of the keyboard to enter their recall, it did occasionally occur. To ensure that our analyses had sufficient data contributed by each participant, we excluded any participants who did not use the retrieval cues at least 12% of the time (i.e., at least 3 times) during either or both blocks. Finally, it was discovered during the experiment that there were two trials in which two of the target words were near-rhymes. Cued recall accuracy for these target words was excluded from analysis for four participants. The trials were corrected for subsequent participants.

Results and Discussion

For all of the experiments, all reported significant results met a criterion of $p < .05$ unless otherwise stated. Measures of effect size (Cohen's *d* or partial eta squared, η_p^2) are reported for all significant *t* or *F* values > 1.

There were no significant differences between cue conditions with regard to accuracy and RTs on the concurrent processing task (*t*s < 1; Table 2). There were also no significant differences between the conditions in terms of recall accuracy, $t(24) = 0.84$, or likelihood of cue use, $t(24) = 0.05$ (Table 2). These null effects were expected as there was no difference between the phonological and semantic blocks up until when participants asked for the cue during recall. Importantly, however, for cued recall accuracy, there was a significant advantage of the semantic over the phonological cues, $t(24) = 3.43$, $d = 0.93$ (Figure 1).

These results show that semantic cues benefit WM recall to a greater extent than phonological cues (i.e., the semantic retrieval cue benefit). This is consistent with previous research demonstrating the influence of semantic processing in verbal WM (Atkins & Reuter-Lorenz, 2008; Loaiza et al., 2011; Shivde & Anderson, 2011), but inconsistent with the original multicomponent view that verbal maintenance is primarily achieved by rehearsal (Baddeley, 1986). This influence of semantic processing, evident in this study by the benefit of administering semantic versus phonological retrieval cues, may be due to the impact of refreshing in a complex span task. As mentioned previously, some work has suggested that the influence of semantic processing may be sensitive to refreshing (Higgins & Johnson, 2013; Rose et al., 2014, 2015). Although our participants responded aloud to the concurrent processing task, which discourages rehearsal and promotes the use of refreshing, there was no manipulation of rehearsal or refreshing in Experiment 1. Thus, Experiment 2 ensured that the semantic retrieval cue benefit was specific to refreshing, and not rehearsing, information in WM by instructing participants to use either rehearsal or refreshing as a maintenance strategy. This experiment also allowed us to test several models' tacit predictions about the efficacy of the cues according to the different maintenance strategies.

Experiment 2

Experiment 1 showed an overall semantic retrieval cue benefit in a complex span task. Experiment 2 was conducted to determine whether this semantic retrieval cue benefit was specific to refreshing as compared to rehearsal. To this end, participants were randomly assigned to one of two groups wherein they were specifically instructed to either rehearse or refresh the memoranda. A strict modality-specific view of verbal WM would predict that phonological cues should benefit recall to a greater extent than semantic cues (Baddeley, 1986). Similarly, the TBRS model would predict that the rehearsal group should maintain memoranda as phonological representations, and thus should exhibit a phonological retrieval cue benefit. However, the TBRS model also espouses a role of refreshing in WM, and thus this model diverges with the former in that it predicts that a semantic advantage should emerge for participants instructed to refresh (i.e., a crossover interaction). The covert retrieval model would similarly predict that the refreshing group should exhibit a semantic retrieval cue benefit due to the influence of LTM, but should be absent or much less evident for participants instructed to rehearse the memoranda because there is no requirement to retrieve the memoranda from LTM. The embedded processes model would also predict a semantic retrieval cue benefit only for the refreshing strategy given the overall effect of semantic processing on refreshing as a search of the central component in WM. Finally, a unitary model of memory would predict an overall semantic retrieval cue benefit regardless of strategies for active maintenance because recall in WM tasks relies on cue-based retrieval from memory (i.e., LTM).

Method

Participants. Fifty-six participants were randomly assigned to either the rehearsal or refreshing instructions condition. Participants were recruited from the University of Fribourg in exchange for partial course credit or cinema ticket coupons. All participants were native French speakers and none had participated in the previous experiment. As in the previous experiment, 12 participants were excluded from the analyses for not using the cues at least 12% of either block. Thus, 44 participants remained for the analysis (22 per instruction condition; 37 female, M_{age} = 22.36, SD = 3.69).

Materials and Procedure. The materials and procedure were identical to Experiment 1, except participants were instructed to either rehearse or refresh the memoranda.

Specifically, participants in the rehearsal group were instructed that they should repetitively rehearse the words in their minds as they would do to remember a telephone number. Conversely, the refreshing group was instructed to "think back" to the words and to not repeat them. This strategy manipulation has been successfully used in a number of other studies (Camos et al., 2011; Johnson, Reeder, Raye, & Mitchell, 2002; Raye et al., 2002). That is, instructing rehearsal or refreshing strategies in participants has yielded different patterns of recall performance (Camos et al., 2011) and activation of distinct frontal areas in the brain (Raye et al., 2002).

After completing both blocks, participants answered a short questionnaire about their instructed strategy at the end of the experiment to serve as a manipulation check. All of the participants in the rehearsal instruction condition correctly identified their strategy, but four of the participants in the refreshing instruction condition misidentified their assigned strategy. Excluding these participants did not change the pattern of results, and thus it was perhaps more likely that the participants were confused about the question rather than that they adopted the incorrect strategy. Furthermore, we also asked about the frequency with which they had used the strategy on a scale of 1 to 5, with 1 labeled as "weakly" and 5 labeled as "all the time." Due to a computer malfunction, three of the participants in the rehearsal condition did not respond to this question. All of the remaining participants responded to this question with at least a 3, and they reported using the strategy with similar frequency between the rehearsal ($M = 3.58$, $SD = 0.90$) and refreshing ($M = 3.64$, $SD = 0.49$) conditions, $t(39) = -0.26$, $p = .798$. There was no significant effect or interaction with cue type of this rating on the semantic cue benefit, Fs < 1.

Design and Scoring. This experiment followed a 2 (strategy: rehearsal, refreshing) x 2 (cue: phonological, semantic) design, with strategy instruction manipulated between-subjects and cue type manipulated within-subjects. The principal dependent variables were recall accuracy, cue use, and cued recall accuracy. As in Experiment 1, it was discovered that there were two trials in which two of the target words were near-rhymes. Cued recall accuracy for these target words was excluded from analysis for 8 participants in the rehearsal group and 6 participants in the refreshing group. The trials were corrected for subsequent participants.

Results and Discussion

Each dependent variable was assessed using a 2 (strategy: rehearsal, refreshing) x 2 (cue: phonological, semantic) mixed analysis of variance (ANOVA). The results are shown in Table 1 and Figure 1.

We first consider performance on the secondary task: processing task accuracy and RTs were similar between the strategy conditions and cue types, Fs < 2.12, ps > .15, except for a significant effect of strategy on task accuracy, $F(1, 42) = 5.57$, $\eta_p^2 = .12$ (Table 2). However, closer inspection of the data shows that both the secondary task performance rehearsal and refreshing strategy conditions was extremely good.

For recall accuracy, the effect of strategy was not significant, $F(1, 42) = 1.68$, $p = .201$, $\eta_p^2 = .04$. Surprisingly, however, there was a significant main effect of cue type, $F(1, 42) = 13.47$, $\eta_p^2 = .24$ and a significant interaction, $F(1, 42) = 5.46$, $\eta_p^2 = .12$. These effects were attributed to the fact that recall was substantially worse for the phonological block than the semantic block for the refreshing group, $F(1, 42) = 18.04$, $\eta_p^2 = .30$, whereas recall did not vary with cue type for the rehearsal group, $F < 1$ (Table 2). The analysis of cue use revealed that this was not necessarily because participants in the refreshing group found the phonological block more difficult. Although the overall effect of cue type was significant, $F(1, 42) = 4.37$, $\eta_p^2 = .09$, there was no difference between the strategy groups in cue use, $F(1, 42) = 1.02$, $p = .319$, $\eta_p^2 = .02$ and no interaction with cue type, $F < 1$ (Table 2). Indeed, as mentioned previously, the semantic and phonological blocks were essentially identical up until the participants asked for a cue, both groups being presented with the same memoranda, and thus there should not be any inherent difference between them in terms of their difficulty.

Given that cue use was similar between the strategy groups, the recall difference must be attributable to errors that some participants made in recall that were unrelated to using the cues. For example, if a participant transposed the order of two memoranda during recall, she would not have asked for cues for either, and yet these would count as recall errors (i.e., serial recall errors). Moreover, although it was strongly discouraged, there were some rare instances where participants did not recall anything nor use the cues, thereby contributing to recall errors (i.e., omission errors). We refer to these as *non-cue related errors*, or errors that were not due to requesting a cue.[1] To address whether the strategy x cue type interaction in recall may be due to these differences in non-cue related errors, we conducted the same analysis on free recall (i.e., scored without regard to serial order) and further excluded any omission errors from the analysis. This analysis showed that the effect of strategy and its former interaction with cue type were not significant, *F*s < 1.27, *p*s > .26, although the overall effect of cue type was smaller but remained, $F(1, 42) = 7.81$, $\eta_p^2 = .16$. This explains the unexpected recall accuracy results. It is not fully understood why serial recall would have been substantially affected in the phonological block for participants instructed to use the refreshing strategy.

The most important analysis concerned cued recall accuracy. The effect of strategy was not significant, $F < 1$, whereas the effect of cue type, $F(1, 42) = 12.45$, $\eta_p^2 = .23$, and the strategy x cue interaction were significant, $F(1, 42) = 4.77$, $\eta_p^2 = .10$. As expected by the TBRS, covert retrieval, and embedded processes models, the locus of this interaction was due

[1] Closer inspection of the data indicated that this was the case: although the two strategy groups did not differ overall in non-cue related errors, $F < 1$, such errors were significantly more likely during the phonological block than the semantic block, $F(1, 42) = 6.63$, $\eta_p^2 = .14$. This effect was qualified by a significant interaction, $F(1, 42) = 4.44$, $\eta_p^2 = .10$: non-cue related errors were similar between the phonological ($M = 0.09$, $SD = 0.07$) and semantic ($M = 0.09$, $SD = 0.08$) blocks for the rehearsal group, $F < 1$, whereas the refreshing group made substantially more non-cue related errors during the phonological block ($M = 0.14$, $SD = 0.11$) than the semantic block ($M = 0.07$, $SD = 0.08$), $F(1, 42) = 10.97$, $\eta_p^2 = .21$.

to the a significant semantic cue benefit in the refresh group, $F(1, 42) = 16.32$, $\eta_p^2 = .28$. However, the rehearsal group did not exhibit any effect of the cues, $F < 1$ (Figure 1), which contradicts the TBRS prediction of a phonological retrieval cue benefit.

The results of this experiment suggest that the previously documented semantic retrieval cue benefit is susceptible to instructions for how to maintain memoranda in WM. Relative to Experiment 1, instructions to use refreshing to maintain memoranda in WM yielded a semantic cue benefit ($d = 1.06$), whereas instructions to rehearse memoranda in WM reduced the semantic cue benefit ($d = 0.25$). Moreover, the size of the semantic retrieval cue benefit was similar between Experiment 1 and the refreshing condition in Experiment 2, as indicated by a non-significant interaction between experiment and cue type, $F < 1$. This supports the prediction that the previously exhibited semantic cue benefit is specific to maintenance of verbal information through attentional refreshing and not with articulatory rehearsal. Furthermore, the results converge with previous evidence that rehearsal and refreshing can be instructed in participants with meaningful differences in the pattern of results (Camos et al., 2011).

The lack of a phonological retrieval cue benefit overall or specifically for the rehearsal group conflicts with the original multicomponent and TBRS models' respective views that rehearsal promotes the phonological representations of the memoranda in WM. Both models predict that rehearsal places memory traces into the phonological loop, and thus their recall should be sensitive to phonological cues. However, the semantic retrieval cue benefit in the refreshing group further conflicts with the original multicomponent model's proposal that there is no attention-based maintenance (Baddeley, 1986; Logie, 2011). This semantic retrieval cue benefit that is exclusive to the refreshing group is more consistent with the covert retrieval model's prediction that refreshing serves to reactivate displaced memory traces from LTM. It is also consistent with the TBRS model that predicts that refreshing uses information from LTM to reconstruct degraded memory traces (Barrouillet & Camos, 2015). Refreshing as a rapid scanning of active representations (Vergauwe & Cowan, 2014, 2015) could also accommodate these findings. Finally, semantic retrieval cue benefit is still consistent with a unitary view of memory. However, the null benefit in the rehearsal group conflicts with the prediction that semantic processing has an overall effect on memory, regardless of maintenance strategies, and thus could not be easily accommodated by the unitary view.

In sum, the overall pattern of findings in Experiment 2 supports the predictions from the covert retrieval model, whereas the TBRS, embedded-processes and unitary models can account for performance observed in the refreshing group. Conversely, the results cannot be accommodated by the strict view that only rehearsal supports verbal maintenance (Baddeley, 1986). Moreover, these results strengthen the suggestion that semantic processing in WM is specific to refreshing rather than rehearsal, thereby providing further insight regarding the nature of the representations that refreshing promotes in WM that are qualitatively distinct from that of rehearsal. To narrow down the possible candidates for how refreshing functions, Experiment 3 manipulated the availability of refreshing as the five compared models make different predictions on how it should impact the semantic retrieval cue benefit.

Experiment 3

Given that Experiment 2 identified that the semantic retrieval cue benefit was specific to refreshing in WM, we examined whether varying attentional refreshing during the task would moderate the benefit in Experiment 3. Previous work has demonstrated that varying the cognitive load of a task specifically impacts the efficiency of attentional refreshing (Barrouillet et al., 2004; Barrouillet et al., 2007; Camos et al., 2009). Thus, Experiment 3 used the same paradigm as in the previous experiments but varied the cognitive load of the concurrent task between-subjects using a typical manipulation of a location versus a parity judgment (e.g., Barrouillet et al., 2007). Three theoretically meaningful results are possible regarding the interaction between cognitive load and cue type. The TBRS model proposes that refreshing uses semantic information from LTM to reconstruct degraded memory traces in WM (Barrouillet & Camos, 2015). This implies a role of semantic processing in WM that is sensitive to the efficacy of refreshing. If refreshing promotes the semantic representation of the memoranda in WM, then increasing the cognitive load of the task should reduce the effectiveness of refreshing, and in turn reduce the semantic retrieval cue benefit. Conversely, it may be that reducing the opportunity to refresh information in WM by increasing the cognitive load increasingly requires episodic LTM to sustain performance (McCabe, 2008; Rose et al., 2014). Thus, the covert retrieval model would predict the opposite interaction to the TBRS, such that increasing the cognitive load of the task should increase the semantic retrieval cue benefit. Another possibility is that the use of refreshing promotes the semantic representation of the memoranda overall in WM, and thus a semantic retrieval cue benefit should be observed regardless of the cognitive load of the task. This would be consistent with a view of refreshing as scanning the central component of WM in the embedded-processes model (Vergauwe & Cowan, 2014, 2015). Finally, the unitary view would also predict an overall semantic retrieval cue benefit regardless of cognitive load due to the presumption that retrieval at the short- and long-term is cue-based, with active maintenance factors playing little role in semantic effects in memory.

Method

Participants. Fifty-two native French speakers were recruited from the University of Fribourg to participate in exchange for partial course credit or cinema ticket coupons. None of the participants had participated in the previous experiments, and they were randomly assigned to the low or high cognitive load conditions. As in the previous experiments, eight participants were excluded from analysis for failing to use the cues at least 12% of the trials in either block. This left 44 participants for the analysis (22 per cognitive load condition; 32 female, M_{age} = 20.59, *SD* = 1.28).

Materials and Procedure. The materials and procedure were identical to Experiment 1, except that the squares used in the processing task were modified so as to include a digit 1-9 within them. The position of the square and digit within the square were randomly selected. In the low load condition, participants were instructed to ignore the digit and respond as to whether the square was up or down on the screen by saying "yes" or "no" out loud, respectively, and also pressing the corresponding keys on the keyboard (i.e., location judgment task). In the high load condition, participants were instructed to ignore the position of the square and respond as to whether the digit was even or odd by saying "yes" or "no" out loud, respectively, and also pressing the corresponding keys on the keyboard (i.e., parity judgment task). We instructed participants to say "yes" or "no" in both conditions so as to equate the responses' syllable length between conditions. Accordingly, whereas the cognitive load of the task was manipulated, the opportunity to engage in articulatory rehearsal was reduced by the concurrent articulation and kept constant across the high and low cognitive load conditions, as implemented in previous studies (Camos et al., 2009; Camos, Lagner, & Loaiza, in press; Camos et al., 2013).

Design. This experiment followed a 2 (cognitive load: low, high) x 2 (cue: phonological, semantic) design, with cognitive load manipulated between-subjects and cue type manipulated within-subjects. As in the previous experiments, the principal dependent variables were recall accuracy, cue use, and cued recall accuracy.

Results and Discussion

Each dependent variable was assessed using a 2 (cognitive load: low, high) x 2 (cue: phonological, semantic) mixed ANOVA. The results are shown in Table 1 and Figure 1.

We first consider the secondary task accuracy and RTs. Although both cognitive load conditions performed the task with sufficient accuracy, as expected, participants in the low cognitive load condition were faster and more accurate than the high cognitive load condition, $F(1, 42) = 131.17$, $\eta_p^2 = .76$ and $F(1, 42) = 16.57$, $\eta_p^2 = .28$, respectively. The effects of cue type and interactions were not significant, Fs < 1.87, ps > .17. These results converge with previous research that parity judgments are more attentionally demanding than location judgments (Barrouillet et al, 2007). Furthermore, they provide an important manipulation check that the conditions were sufficiently different in their cognitive load.

For recall accuracy, as expected, there was a significant main effect of cognitive load, $F(1, 42) = 9.79$, $\eta_p^2 = .19$, such that the participants in the low load condition recalled significantly more words than the high load condition. The effects of cue type and the interaction were not significant, Fs < 1.01, ps > .31. Regarding the likelihood of cue use, there was again a significant main effect of cognitive load, $F(1, 42) = 7.24$, $\eta_p^2 = .15$, such that those in the low load condition were significantly less likely to ask for a cue than those in the high load condition. The effect of cue type ($F < 1$) and the interaction ($F(1, 42) = 2.41$, $p = .128$, $\eta_p^2 = .05$) were not significant.

Finally, for cued recall accuracy, there was a significant effect of cue type, $F(1, 42) = 8.96$, $\eta_p^2 = .18$, such that semantic cues were more effective for cued recall than the phonological cues, replicating the semantic retrieval cue benefit of the previous experiments. The effect of cognitive load was not significant, $F < 1$. Finally, and most importantly, there was no significant interaction, $F < 1$. As many researchers have increasingly pointed out, p values do not provide evidence in favor of the null hypothesis, and thus it is not clear whether the observed null interaction is due to a failure to observe a true effect. However, given the aforementioned predictions of the various tested models, it is important to identify the evidence for the null interaction. Bayesian inferential statistics provide a means to assess the relative evidence for a given model (e.g., a null interaction model) over another model (e.g., a model positing an interaction). Accordingly, we computed the Bayesian Information Criteria (BIC) for this effect, $p_{BIC}(H0|D) = .87$. A probability above .75 is considered positive evidence for the null hypothesis (Masson, 2011), thus supporting the absence of interaction between cognitive load and cue type. We also considered whether the semantic retrieval cue benefit was similar between this experiment and Experiment 1 that had the same conditions as the low load condition in the current experiment. Although the semantic retrieval cue benefits were nominally smaller in the low cognitive load ($d = 0.59$) and high cognitive load ($d = 0.57$) conditions of Experiment 3 relative to that of Experiment 1 ($d = 0.93$), the interaction between experiment and cue type was not significant, $F < 1$. Thus, the semantic retrieval cue benefit was consistent across cognitive load.

These results further demonstrate the importance of semantic processing in WM, consistent with previous research (Loaiza et al., 2011; Rose et al., 2014; Shivde & Anderson, 2011). However, the current experiment expands upon this research on the role of semantic processing by examining whether the identified specificity of the semantic retrieval cue benefit to refreshing (Experiment 2) can further adjudicate between different WM models' conceptions of refreshing. Whereas the TBRS and covert retrieval model predicted an interaction between cognitive load and the semantic retrieval cue benefit, the positive evidence for a null interaction supports the embedded processes model's recent proposal regarding how refreshing in WM operates and the unitary memory view. Specifically in the embedded processes model, the results comport with the view that refreshing emphasizes semantic representations overall, regardless of cognitive load, as a consequence of functioning as a rapid scanning of active representations that are embedded within the broader context of LTM (Vergauwe & Cowan, 2014, 2015). Alternatively, the semantic retrieval cue benefit can reflect retrieval from episodic LTM as proposed by the unitary view of memory because WM is not differentiated from LTM, and recall results from cue-based retrieval from memory.

However, the unitary memory view is not the only one that may conceive the observed semantic retrieval cue benefit as reflecting episodic LTM retrieval. Indeed, in Experiments 1 to 3, cues were administered only if participants required help to recall the memoranda, and thus only when information was not recovered, and thus presumably lost, from WM. This may have prompted a search of episodic LTM. This alternative account is congruent with all the WM models we contrasted in this study (i.e., TBRS, covert retrieval and embedded processes models). One method to examine whether the semantic retrieval cue benefit observed in the paradigm used in Experiments 1 to 3 reflects retrieval from episodic LTM or not is to present cues for all memoranda and to consider the intention to maintain the memoranda in WM.

Intention to learn information has historically been used in episodic LTM paradigms (e.g., Craik & Tulving, 1975; Hyde & Jenkins, 1973) by instructing participants to learn the memoranda for an upcoming test (intentional encoding) or having the memory test as a complete surprise (incidental encoding). Encoding is rarely manipulated this way in WM paradigms (except see Rose & Craik, 2012), most probably because keeping the surprise aspect of the incidental encoding condition is necessarily constrained after one trial; participants would expect a memory test for further trials thereafter regardless of instruction. Such a manipulation in a WM paradigm would shed light on whether the semantic retrieval cue benefit observed in Experiments 1 to 3 relies on episodic LTM retrieval, but would also allow us to test the contrasted predictions issued from the unitary memory view and the WM models.

According to the unitary memory view, recall is always based on a cue-based search in LTM. As a consequence, although recall performance should benefit from intentional encoding, semantic cues should have a strong beneficial effect on recall whether the encoding condition is intentional or incidental, because they provide better cues to retrieve information in LTM relative to phonological cues. However, if the semantic retrieval cue benefit observed in the previous experiments results from the loss of WM traces and the search of episodic LTM, it should only occur in the incidental condition, in which participants did not maintain information in WM and can only rely on episodic traces to respond. Alternatively, if the semantic retrieval cues benefit reflects the role of refreshing in promoting semantic representations in WM, it should be larger under intentional compared to incidental encoding. These hypotheses were tested in Experiment 4.

Experiment 4

The design of Experiment 4 was very similar to Experiment 3 in that we examined whether the efficacy of semantic over phonological cues varied as a function of cognitive load. However, there were three key differences. First, to address whether cued recall accuracy in the previous experiments was primarily due to retrieval from episodic LTM given that these are forgotten items in WM, all of the cues were presented on the screen for each item without the participants asking for them. Second, the previous experiments all used intentional encoding wherein the participants were explicitly instructed to remember the memoranda and expected to be tested on them. We manipulated an additional variable of the expectation of a memory test (i.e., type of encoding, incidental versus intentional) to examine the contribution of episodic LTM retrieval to the semantic retrieval cue benefit. Accordingly, half of the participants did not expect that their memory for the presented memoranda would be tested. Thus, the participants should not actively maintain the memoranda in WM and their performance would presumably be driven exclusively by retrieval from episodic LTM (Rose & Craik, 2012). Finally, the surprise recall aspect necessarily required that only one trial per participant could be collected because further trials in an incidental encoding condition would likely be contaminated with the expectation of a memory test thereafter. Moreover, only one trial per condition requires that the design is fully between-subjects. Given these constraints and the large number of participants that would be required to overcome them, we opted to use Amazon Mechanical Turk (AMT) for data collection for Experiment 4. AMT is an increasingly utilized tool for collecting reliable experimental data from a large number of participants (Crump, McDonnell, & Gureckis, 2013; Mason & Suri, 2012). We were thus able to collect one trial from a very large sample of adults with a similar age and educational background as the participants in the previous experiments.

This experiment was designed to test the previous hypotheses regarding the semantic retrieval cue benefit reported in the previous experiments (Table 1). It may be the case that benefit of semantic retrieval cues may not interact with intention to learn, thereby supporting the unitary memory view. Alternatively, an interaction between cue type and encoding condition with a semantic retrieval cue benefit appearing only in the incidental condition would indicate that this benefit results from retrieval from episodic LTM. Finally, if the semantic cue benefit is larger under intentional compared to incidental encoding, then this would support the role of refreshing in promoting semantic representations in WM, above and beyond the overall benefit semantic cues would be expected to provide to memory. Furthermore, how this semantic retrieval cue benefit interacts with cognitive load within the intentional encoding condition would differently support the predictions of the three aforementioned accounts of how refreshing functions as investigated in Experiment 3.

Method

Participants and Design. This experiment utilized 2 (encoding: incidental, intentional) x 2 (cognitive load: low, high) x 2 (cue type: semantic, phonological) fully between-subjects design. The main dependent variable was recall, and we also assessed RTs and accuracy on the secondary task.

Due to the atypical nature of the experiment having a fully between-subjects design and only one trial per participant, we calculated an a priori power analysis to determine the required number of participants to detect an effect of size $f = 0.25$ with a power level $(1 - \beta)$ of 0.80. This yielded a total sample size of 128 participants, and we aimed for this sample size with an approximately even representation across the eight experimental groups (total *n* per group: incidental-low-semantic $n = 16$; incidental-low-phonological $n = 16$; intentional-low-semantic $n = 16$; intentional-low-phonological $n = 17$; incidental-high-semantic $n = 17$; incidental-high-phonological $n = 16$; intentional-high-semantic $n = 18$; intentional-high-phonological $n = 16$). Due to the counterbalancing of the memoranda (see Materials and Procedure), there were 16 groups in total, and participants were randomly assigned to one of the groups.

In total, 244 people were recruited, but only 132 participants were included in the final analysis (74 female, $M_{age} = 29.03$, $SD = 4.41$, range = 18 – 35). The reasons for exclusion were: quitting the program during the middle of the experiment (usually the practice phase; $n = 62$); a mismatch in the memory instructions and expectations (e.g., participants reporting that they expected a memory test for the memoranda when they shouldn't have; $n = 12$); failing to follow instructions (i.e., participants reporting that they did not read the memoranda aloud or solve the secondary task aloud; $n = 23$); or other reasons (e.g., entering digits 1-5 in the spaces provided during recall, completing the study twice; $n = 15$). Participants were recruited via AMT and compensated $0.50 for approximately 5 min of their time. AMT worker requirements ensured that participants were currently located in the United States, had a US high school education and had a human intelligence task (HIT) approval rate of greater than 90%. Participants also reported being native English speakers.

Materials and Procedure. The initial qualification survey was administered in Qualtrics and the experiment was programmed in Inquisit and administered via participants' web browsers with a downloaded plugin. Two sets of five memoranda were drawn from the English word triads reported in Rose et al. (2010) so that they had a similar overall forward associative strength ($M = 0.34$, $SD = 0.07$, range = 0.20 – 0.46) as the previous experiments' French memoranda. The associative strength was similar between two sets of words (set A: $M = 0.35$, $SD = 0.08$; set B: 0.32, $SD = 0.08$, $t(8) = 0.69$, $p = .51$). The sets were also checked to ensure that the words were not semantically related or rhymed within each set. These sets were counterbalanced across participants and their order of presentation was randomized within the lists for each participant.

AMT workers with the aforementioned prescreened requirements were invited to complete our HIT entitled "respond to speeded mental tasks" in which they were told that they would respond as quickly and accurately as they could to presented stimuli. No mention of the memory element of the task was specified until later on in the experiment. The workers were warned that they should accept the HIT only if they had never done the study before, were using a desktop or laptop computer, and were working in a quiet place with no distractions or other ongoing activities. After clicking the link provided in the HIT, the workers were redirected to a Qualtrics survey that first determined whether they qualified for the experiment. If workers reported that they were aged 18 – 35, native English speakers, and the survey determined they were not using a mobile device, they were then redirected to the consent form. The other workers who did not meet the experiment criteria were redirected to an end of survey page that informed them they did not meet the qualifications. After reading the consent form and agreeing to participate, the participants were then redirected to one of the Inquisit links that presented the experimental condition to which they were randomly assigned.

The rest of the procedure for the experiment was very similar to Experiment 3. Participants first completed a practice phase in which they saw 20 squares with digits at their centers successively appear on the screen one at a time for 700 ms (300 ms ISI) each. Depending on the condition, participants were instructed to respond as to whether the square was in the upper half of the screen or not (low cognitive load) or whether the digit was even or not (high cognitive load). Participants responding using a designated right- or left-hand key and were also instructed to respond “yes” or “no” aloud for each trial. The practice phase was repeated until they achieved an 85% criterion. After finishing the practice phase, they received instructions for the second, critical phase of the experiment. They again saw squares with digits successively appear on the screen and responded as they had during the practice phase. They were also instructed that they would see words presented in red font for 1000 ms (500 ms ISI) that they should read aloud in-between the presentation of and responding to the squares/digits. Depending on the encoding condition, the participants received different instructions regarding the words: either that words were meant to distract them from their main task of responding to the squares/digits, and so to focus more on squares/digits task (incidental encoding) or they were explicitly told to try their best to remember the words for an upcoming memory test, but to try their best to respond to the secondary task as well (intentional encoding).

At the end of the trial, participants saw a screen with an invisible 2 x 5 grid with a cue word next to each space to enter the recalled words. Participants received instructions at the top of the screen to try to recall the words in the order of their presentation, and that the cues were clues to help them to try to recall the words. Depending on the condition, the cue words either were phonologically or semantically related to the memoranda, and this was specified to the participants in the instructions. Participants typed their responses and pressed enter to move onto the next word to recall. After they finished the recall, a questionnaire of three yes/no questions was presented that asked participants about their performance on the previous task. Specifically, they were asked whether they said yes or no aloud while pressing the keys during the secondary task, whether they read the words aloud only once when they appeared, and whether they expected their memory would be tested for the presented words. Participants' responses to these questions qualified them for inclusion in the analyses as specified previously. Following the questionnaire, participants were presented with a debriefing and then entered their AMT identification for their compensation.

Results and Discussion

Each dependent variable was assessed using a 2 (encoding: incidental, intentional) x 2 (cognitive load: low, high) x 2 (cue: phonological, semantic) independent measures ANOVA. The results are shown in Table 1 and Figure 2.

The first analysis concerning performance on the secondary task yielded a significant effect of cognitive load for accuracy, $F(1, 124) = 18.12$, $\eta_p^2 = .13$, and RTs, $F(1, 124) = 317.66$, $\eta_p^2 = .72$. These effects are consistent with the previous experiment and other laboratory-based findings that increasing the cognitive load impairs performance on the secondary task. There was also a marginally significant effect of encoding on RTs, $F(1, 124) = 3.92$, $p = .050$, $\eta_p^2 = .03$, such that participants in the intentional condition (M = 493 ms, SD = 81) were slightly slower to respond than those in the incidental condition (M = 478, SD = 86). This is plausible given that participants in the intentional encoding condition were explicitly told to remember the words, and thus their active intention to maintain the memoranda slightly slowed their responses to the secondary task (see Vergauwe, Camos, & Barrouillet, 2014, for similar findings). All other effects and interactions for accuracy and RTs were not significant, Fs < 1.54, ps > .21 (Table 2).

Given that the participants were presented with the cues without asking for them as in the previous experiments, there are no measures of recall accuracy or cue use, and thus the remaining analysis concerned cued recall accuracy (Figure 2). As expected, participants in the intentional group showed greater recall (M = 0.70, SD = 0.31) than those in the incidental group (*M* = 0.47, *SD* = 0.35), $F(1, 124) = 19.65$, $\eta_p^2 = .14$. Thus, as is consistent with much prior research, advance warning of a memory test improved performance. Recall also significantly improved when semantic cues were presented (*M* = 0.71, *SD* = 0.31) compared to phonological cues (*M* = 0.46, *SD* = 0.34), $F(1, 124) = 22.13$, $\eta_p^2 = .15$. Importantly, all other effects and interactions were not significant, *F*s < 2.19, *ps* > .14. We focus in more detail on the most relevant interactions that speak to the various predictions for the experiment.

First, the cue x cognitive load interaction was not significant, $F < 1$, with the evidence positively supporting a null effect, $p_{BIC}(H0|D) = .91$. This replicates Experiment 3's findings showing that the semantic retrieval cue benefit did not change as a function of cognitive load. Secondly, the encoding x cue interaction was also not significant, $F(1, 124) = 2.18$, $p = .142$, $p_{BIC}(H0|D) = 0.78$. This indicates that whether the memory test was a surprise or not had no impact on the semantic retrieval cue benefit, and thus the source of the effect in the previous experiments is not simply driven by forgotten information from WM. It is further important to note that the interaction between encoding and cognitive load was also not significant, $F(1, 132) = 1.55$, $p = .215$, $p_{BIC}(H0|D) = .84$. Thus, the null impact of cognitive load was consistent regardless of whether the encoding was intentional or incidental. Finally, the three-way interaction between encoding, cognitive load, and cue type was not significant, $F < 1$, $p_{BIC}(H0|D) = .92$. Thus, there was substantial evidence that the semantic retrieval cue benefit was consistent across cognitive load, thus contradicting the TBRS and covert retrieval models, but consistent with the embedded processes model. However, the benefit was also consistent across intention to learn, negating the predictions from the TBRS, covert retrieval, and embedded processes models.

In summary, these results replicate and extend the results of Experiment 3. Even when the cues were presented for all of the memoranda, there was still an overall semantic retrieval cue benefit that did not interact with cognitive load or intention to learn. Moreover, just as in Experiment 3, cognitive load did not impact cued recall accuracy. The positive evidence in favor of the null interaction between the semantic retrieval cue benefit and the intention to learn contrasted with the predictions from the WM models that emphasize the role of refreshing to promote semantic representations in verbal WM. Overall, these results are most consistent with a unitary view of memory that does not posit a distinction between WM and episodic LTM, but instead comprises a cue-based search and retrieval of memory (Crowder, 1982; Melton, 1963; Nairne, 2002).

General Discussion

The present series of experiments explored attentional refreshing in verbal WM, especially with the respect to following three goals: (1) examine the nature of representations that refreshing emphasizes during a complex span task, (2) narrow down the several candidate proposals that have been advanced in the literature, thereby elucidating how refreshing functions to maintain memoranda in WM, and (3) enlighten the relationship between WM and LTM. To address these goals, we employed a novel paradigm in a series of experiments that provided semantic and phonological cues during recall from a complex span task. The results collectively showed a semantic retrieval cue benefit, such that semantic retrieval cues were more beneficial to WM recall over phonological cues. This converges with a growing literature suggesting that semantic processing is important in WM (Atkins & Reuter-Lorenz, 2008; Loaiza et al., 2011; Rose et al., 2014; Rose et al., 2015; Shivde & Anderson, 2011), and is incongruent with a strict view of verbal maintenance as exclusively phonological as in the original multicomponent model (Baddeley, 1986; see also Baddeley & Logie, 1999; Logie, 2011). We investigated whether semantic processing is evident in WM because refreshing may emphasize the semantic characteristics of memoranda (Higgins & Johnson, 2013). Although Experiment 2 showed that this semantic retrieval cue benefit was specific to refreshing rather than rehearsal by varying the maintenance strategy instructions given to participants, Experiments 3 and 4 indicated that the semantic retrieval cue benefit occurred regardless of manipulations that were intended to vary the active maintenance in WM (by varying either the cognitive load or the intention to learn). Thus, semantic processing effects in verbal WM may occur because of the underlying cue-based search in episodic LTM that contributes to memory more generally and regardless of active maintenance. In sum, this study makes a novel contribution to this literature by highlighting the influence of semantic representations in WM and by allowing adjudication between different theoretical conceptions of refreshing. It has strong implications for the ongoing debate regarding the distinction between WM and LTM. We will discuss these three issues in turn.

Semantic representations in WM

The first goal of the current study was to examine the nature of representations that are promoted as a function of refreshing in WM. Previous research has shown that refreshing and rehearsal are distinct maintenance mechanisms, evident in their contribution to WM recall (Camos et al., 2009), episodic LTM (Camos & Portrat, 2015; Loaiza & McCabe, 2013; Raye et al., 2002; Rose et al., 2014), and underlying neural substrates (Raye et al., 2007; Raye et al., 2002). Further studies have qualified this distinction by investigating the nature of the processes underlying WM maintenance as a function of either mechanism. Congruent with prior research on immediate recall from STM (Baddeley, 1966; Baddeley et al., 1975), Camos and colleagues (2011; 2013; Mora & Camos, 2013) have shown that rehearsal strongly emphasizes the phonological characteristics of memoranda in complex span tasks. Similar to previous WM models (Baddeley, 1986), Camos and colleagues argued that this is due to the nature of rehearsal as a peripheral, domain-specific mechanism that regenerates the phonological representations of memoranda in WM. Accordingly, rehearsal is specific to maintenance of verbal information and can be prevented by concurrent articulation. Conversely, refreshing is not dependent on the phonological characteristics of the memoranda. This was evident in the disappearance of the phonological similarity effect when refreshing was instructed for use during a WM task (Camos et al., 2011). Thus, refreshing was argued to operate independently of the phonological characteristics that constrain rehearsal. The question remained regarding which characteristics of memoranda are relevant (or at least more relevant) to refreshing. Previous studies have suggested that semantic processing evident in verbal WM (e.g., Loaiza et al., 2011; Rose et al., 2014; Rose & Craik, 2012) may be moderated by attention-based factors that are independent from phonological effects (Atkins & Reuter-Lorenz, 2008; Shivde & Anderson, 2011). For example, Nishiyama (2014) showed that manipulating the attentional demand of a concurrent task (e.g., tapping) more significantly reduced performance on a synonym (semantic) recognition task relative to performing the task with articulatory suppression. Rose and colleagues (2014, 2015) also showed that increasing the attentional demand of a secondary task increased the levels-of-processing effect in WM recall. Thus, these studies indicated a link between attention-based maintenance (i.e., refreshing) and semantic representations in WM.

The results of this study also collectively indicated that semantic representations are important in WM. Experiment 2 demonstrated that the semantic retrieval cue benefit was eliminated when participants were instructed to use rehearsal rather than refreshing as a maintenance strategy. This result comports with a growing literature distinguishing rehearsal and refreshing (see Camos, 2015, 2017, for review) given that rehearsal seems to reduce the importance of semantic processing in WM compared to refreshing (Rose et al., 2014, 2015). For example, Rose and colleagues showed that participants instructed to rehearse memoranda during a retention interval did not exhibit any recall difference between shallow and deep, semantic processing at encoding. Rose and Craik (2012) also showed a benefit of semantic processing during encoding only when active maintenance processes were eliminated in an incidental encoding condition using a WM paradigm. However, the semantic retrieval cue benefit in the present study remained regardless of factors intended to manipulate refreshing specifically (i.e., cognitive load; Experiments 3 and 4) or active maintenance in WM (i.e., intention to learn; Experiment 4). This last finding is in line with previous episodic LTM studies that have shown that varying the intention to learn does not moderate the benefit of semantic processing during encoding (Craik & Tulving, 1975; Hyde & Jenkins, 1973). This suggests that the previously documented evidence of semantic processing in WM may have less to do with the impact of domain-general, attention-based maintenance via refreshing and more to do with a cue-based search of a unitary memory.

Although the results were consistent overall with the predictions from a unitary view of memory, the lack of semantic retrieval cue benefit in the rehearsal condition in Experiment 2 diverged from the tacit prediction of this view that semantic processing should be evident regardless of active maintenance strategies in WM. This finding suggests that there may be other possible interpretations of the current results. For example, another interpretation may be that providing cues during recall increased the reliance on episodic LTM and thereby reduced the potential impact of active maintenance factors in WM. Indeed, there is a substantial amount of research showing that WM task conditions do modulate the way memoranda are encoded and retained in WM with consequences for how they are represented and retrieved from episodic LTM (e.g., Jacoby & Bartz, 1972; Loaiza et al., 2011; Mazuryk & Lockhart, 1974; Rose et al., 2014; Rose & Craik, 2012). Furthermore, as detailed in the Introduction, much research has suggested that varying the attentional demand of a concurrent task also greatly affects recall from WM (Barrouillet et al., 2007; Barrouillet et al., 2011) and even episodic LTM (Camos & Portrat, 2015). Thus, it is somewhat surprising that the manipulation of cognitive load in Experiments 3 and 4 did not affect cued recall accuracy overall, much less the semantic retrieval cue benefit. This result is even more intriguing when one considers the fact that the classic cognitive load effect was demonstrated in accurate recall in Experiment 3, but it was not shown in cued recall for the same experiment. This is inconsistent with the aforementioned studies supporting the notion that active maintenance and encoding processes in WM affect retrieval. However, all of these studies (e.g., Barrouillet et al., 2007; Camos & Portrat, 2015) used recall rather than other kinds of retrieval methods, such as cued recall, as we used here. Thus, the semantic retrieval cue benefit demonstrated here may be more indicative of episodic LTM resources, but may also have superseded the impact of most active maintenance manipulations (except rehearsal). This is consistent with Atkins and Reuter-Lorenz (2008), who showed that the effect of attentional demand on semantic interference was evident during recall but not recognition of memoranda in WM. This suggests that providing information (e.g., cues in cued recall or probes in recognition tasks) may increase the contribution of episodic LTM to such an extent that it overshadows most factors that manipulate the memoranda's representation in WM. To summarize, the current findings inform our understanding of how refreshing functions in that its influence is separable to the contribution of episodic LTM to immediate recall from WM, thereby contradicting the overviewed WM models' predictions regarding how refreshing operates on semantic representations.

Given this powerful impact of their use during retrieval, it is interesting to consider what makes semantic cues more beneficial than phonological cues. One possibility is that semantic cues may have fewer candidate targets compared to phonological cues that may have more neighbors that compete with the target. Unfortunately, it is difficult to equate semantic and phonological cues on such a pertinent level. However, this possibility resonates with the aforementioned issue that providing cues may encourage reconstructive memory resources underlying episodic LTM. That is, by providing cues during retrieval even at short-term intervals, factors that are known to affect cue-based search in episodic LTM (such as cue overload, Watkins & Watkins, 1975) may dominate over active maintenance manipulations. Another possibility is that the memoranda each have dissociable semantic and phonological traces in WM. Perhaps these traces deteriorate (whether by decay or interference) at different rates, such that the phonological trace deteriorates faster than the semantic trace. This would also yield a semantic retrieval cue benefit that is reduced when participants are encouraged to use rehearsal and thereby emphasize the phonological trace of the memoranda and resulting in null semantic retrieval cue benefit. Further research would be necessary to disentangle these possibilities, but overall the results of the present study are consistent with the view that the semantic retrieval cue benefit indicates a strong contribution of episodic LTM to what is typically considered WM (Unsworth & Engle, 2007).

Implications for refreshing and the overlap between working memory and long-term memory

In addition to highlighting the impact of semantic processing in verbal WM, the results of the current study address several different models concerning how refreshing and active maintenance in WM may operate. Our results also have implications for the overlap between WM and LTM, especially given that refreshing interfaces with LTM in many of the models' predictions. The original multicomponent model of WM predicted that verbal maintenance is primarily achieved by a domain-specific rehearsal, whereas attention has no role in verbal maintenance (Baddeley, 1986; Logie, 2011). This would mean that refreshing technically should not even exist (although, see later discussion of the updated multicomponent model, Baddeley, 2000, 2012). Conversely, refreshing serves as a defining feature of the TBRS model (Barrouillet et al., 2004; 2007), such that refreshing supports maintenance in WM by reconstructing degraded memory traces in the episodic buffer using information from LTM (Barrouillet & Camos, 2015). Refreshing is also completely independent from rehearsal, according to the TBRS model, such that rehearsal maintains information in a phonological format within the phonological loop, and is accordingly constrained by articulatory suppression. The efficacy of refreshing, on the other hand, is impeded upon by increasing the cognitive load of a concurrent task. The third reviewed proposal of refreshing from the covert retrieval model (McCabe, 2008) suggests that refreshing serves as a covert retrieval of less active, displaced information from LTM back into primary memory. Accordingly, when refreshing is impeded upon by a high cognitive load, WM recall is increasingly dependent on LTM (see Rose et al., 2014; Rose et al., 2015 for similar prediction). For all three of these models, WM and LTM represent separable subsystems of memory (but see Loaiza et al., 2015a; Loaiza & McCabe, 2012). We further considered a proposal of refreshing originating from the embedded processes model (Cowan, 1999), which views WM as an activated subset of LTM. The embedded processes model distinguishes between different hierarchical states of activation: within LTM, a subset of its representations is strongly activated and accessible (i.e., activated LTM), and the focus of attention contains the few chunks of information most immediately accessible to conscious awareness (for a similar proposal, see the concentric model by Oberauer, 2002). Although this model originally described refreshing as a refocalization of attention on previously activated representations, it has been more recently suggested that refreshing acts as a rapid scanning of the central component of WM (Vergauwe & Cowan, 2015). Finally, we also considered the possibility that refreshing or active maintenance more generally is not necessary to explain any semantic effects in WM that instead reflect some cue-based search in episodic LTM (e.g., Nairne, 2002). This unitary view makes no distinction between WM and LTM; instead, they comprise the same underlying memory system. The current study allowed for a test of each of these models by examining the occurrence of a semantic retrieval cue benefit for information forgotten during recall from a complex span task and for intentionally vs. incidentally memorized information.

Overall, the results of the current study collectively support the unitary view of memory, such that the semantic retrieval cue benefit in WM recall reflects the contribution of a cue-based search of episodic LTM. This view was most strongly supported by the findings that the semantic retrieval cue benefit was consistent across manipulations of cognitive load (i.e., Experiments 3 and 4) and intention to learn (i.e., Experiment 4). This has strong implications regarding the nature of refreshing and relationship between WM and LTM, as there may be no role of refreshing in emphasizing semantic representations in WM, and that WM and LTM are not separable systems. However, as mentioned previously, the finding that rehearsal nullified the semantic retrieval cue benefit in Experiment 2 is difficult for the unitary view to accommodate. We previously discussed the possibility that the phonological and semantic traces of the memoranda deteriorate at different rates, and rehearsal counteracts this by emphasizing the phonology of the memoranda. This still implies a role of rehearsal, however, which is typically denied by proponents of the unitary view (e.g., Nairne, 2002). The finding that cognitive load does impact recall from WM in Experiment 3, corroborating extensive previous literature (e.g., Barrouillet et al., 2011), is also inconsistent with the unitary view that active maintenance in WM is not important to what is ultimately a cue-based search of episodic LTM. It is not immediately clear how traditional unitary views may accommodate cognitive load effects. Thus, the current results overall were most consistent with the unitary view, but not perfectly. Some additional assumptions are necessary for it to fully account for these findings.

Alternatively, that the semantic retrieval cue benefit did not vary with cognitive load or intention to learn the memoranda (both affecting active maintenance) may instead suggest that this benefit relies on what happens during the encoding of memoranda, before any maintenance activity. This implies that existing semantic networks (i.e., semantic LTM) facilitate the encoding of the verbal memoranda as semantic representations. The only situation in which this is not the case is when using rehearsal to maintain the memoranda that instead emphasizes their phonological characteristics (e.g., Camos et al., 2011), thereby reducing the efficacy of semantic retrieval cues. This alternative explanation further implies that refreshing still may have a role to sustain memory traces overall, but this role does not comprise operating on the semantic representations of the memoranda as predicted by many of the WM models (Table 1). In other words, refreshing does not modulate the status of the semantic representations that were already preferentially and stably encoded as such. This is consistent with Oberauer's (2009) finding that the beneficial effect of semantically similar distracters on WM recall did not vary with cognitive load. This would also converge with the frequent finding that cognitive load impacts recall from WM (as in Experiment 3; Barrouillet & Camos, 2015), but not cued recall that emphasizes the different characteristics of the memoranda that were encoded during their initial presentation. Thus, refreshing may have nothing to do with semantic LTM. Rather, once the semantic characteristics of the memory traces are encoded, refreshing reactivates these traces without support from LTM. This idea conflicts with other studies suggesting that refreshing interacts with LTM factors like the semantic characteristics or the level of processing engaged during encoding of the memoranda (e.g., Higgins & Johnson, 2013; Loaiza et al., 2015a; Rose et al., 2014, 2015). Future research will be necessary to disentangle the conditions in which LTM factors may contribute beyond their influence during encoding.

Taken together, the consistent semantic retrieval cue benefit across variations of cognitive load and intention to learn suggests that the effect may reflect the facilitation of LTM on encoding that is not further moderated by refreshing. Furthermore, as mentioned previously, the absence of an overall effect of cognitive load on cued recall also indicates that the contribution of LTM resources when using cued recall may overwhelm the effect of active maintenance via refreshing on recall. These results collectively provide insight that (1) LTM strongly contributes to performance in WM tasks, and (2) refreshing functions by reactivating memory traces independently from the influence of LTM.

Conclusions

In summary, the current series of experiments demonstrated a semantic retrieval cue benefit for memoranda studied during WM that was only eliminated for a rehearsal-based maintenance strategy. Neither manipulating refreshing via the cognitive load of the secondary task nor the intention to learn the memoranda moderated the semantic retrieval cue benefit. This suggests that, whereas rehearsal may promote the phonological characteristics of memoranda in WM (Camos et al., 2013, 2011), previously documented semantic effects may originate from the contribution of episodic LTM to WM performance, regardless of refreshing or even intention to learn the memoranda. This study expands the developing literature regarding the role of semantic processing in short-term retention (Martin, 2005; Shivde & Anderson, 2011), and the results speak to different models' conceptions of refreshing. While this study makes it further clear that verbal maintenance in WM is not exclusively phonological (Baddeley, 1986), the results also suggest refreshing does not vary the status of semantic representations in WM. Instead, semantic effects in WM may reflect the contribution of episodic LTM even for short-term retention.

References

Atkins, A. S., & Reuter-Lorenz, P. A. (2008). False working memories? Semantic distortion in a mere 4 seconds. *Memory & Cognition, 36*, 74-81. doi:10.3758/Mc.36.1.74

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362-365. doi:10.1080/14640746608400055

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

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). New York: Cambridge University Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and structure of short-term-memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589. doi:10.1016/S0022-5371(75)80045-4

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. doi:10.1037/0096-3445.133.1.83

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. doi:10.1037/0278-7393.33.3.570

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review, 118*(2), 175-192. doi:10.1037/a0022324

Camos, V. (2015). Storing verbal information in working memory. *Current Directions in Psychological Science, 24*(6), 440-445. doi:10.1177/0963721415606630

Camos, V. (2017). Domain-specific vs. domain-general maintenance in working memory: Reconciliation within the time-based resource sharing model. In B. Ross (Ed.), *The Psychology of Learning and Motivation.* (Vol. 67).

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. doi:10.1016/j.jml.2009.06.002

Camos, V., Lagner, P., & Loaiza, V. M. (in press). Maintenance of item and order information in verbal working memory. *Memory.* doi:10.1080/09658211.2016.1237654

Camos, V., Mora, G., & Barrouillet, P. (2013). Phonological similarity effect in complex span task. *Quarterly Journal of Experimental Psychology, 66*(10), 1927-1950. doi:10.1080/17470218.2013.768275

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. doi:10.3758/s13421-010-0011-x

Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin & Review, 22*(4), 1029-1034. doi:10.3758/s13423-014-0772-5

Campoy, G., Castella, J., Provencio, V., Hitch, G. J., & Baddeley, A. D. (2015). Automatic semantic encoding in verbal short-term memory: Evidence from the concreteness effect. *Quarterly Journal of Experimental Psychology, 68*(4), 759-778. doi:10.1080/17470218.2014.966248

Conrad, R. (1963). Acoustic confusions and memory span for words. *Nature, 197*, 1029-1030.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55*, 75-84. doi:10.1111/j.2044-8295.1964.tb00899.x

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429-432. doi:10.1111/j.2044-8295.1964.tb00928.x

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). New York: Cambridge University Press.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294. doi:10.1037//0096-3445.104.3.268

Crowder, R. G. (1982). The demise of short-term-memory. *Acta Psychologica, 50*(3), 291-323. doi:10.1016/0001-6918(82)90044-0

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *Plos One, 8*(3). doi:ARTN e5741010.1371/journal.pone.0057410

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual-Differences in Working Memory and Comprehension - a Test of 4 Hypotheses. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*(5), 972-992. doi:10.1037/0278-7393.18.5.972

Ferrand, L., & Alario, F. X. (1998). Word association norms for 366 names of objects. *Annee Psychologique, 98*(4), 659-709.

Flegal, K. E., Atkins, A. S., & Reuter-Lorenz, P. A. (2010). False memories seconds later: The rapid and compelling onset of illusory recognition. *Journal of Experimental Psychology: Learning Memory and Cognition, 36*(5), 1331-1338. doi:10.1037/a0019903

Flegal, K. E., & Reuter-Lorenz, P. A. (2014). Get the gist? The effects of processing depth on false recognition in short-term and long-term memory. *Memory & Cognition, 42*(5), 701-711. doi:10.3758/s13421-013-0391-9

Hamilton, A. C., Martin, R. C., & Burton, P. C. (2009). Converging functional magnetic resonance imaging evidence for a role of the left inferior frontal lobe in semantic retention during language comprehension. *Cognitive Neuropsychology, 26*(8), 685-704. doi:10.1080/02643291003665688

Hanten, G., & Martin, R. C. (2000). Contributions of phonological and semantic short-term memory to sentence processing: Evidence from two cases of closed head injury in children. *Journal of Memory and Language, 43*(2), 335-361. doi:10.1006/jmla.2000.2731

Higgins, J. A., & Johnson, M. K. (2013). Lost thoughts: Implicit semantic interference impairs reflective access to currently active information. *Journal of Experimental Psychology: General, 142*(1), 6-11. doi:10.1037/a0028191

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language, 46*(1), 85-98. doi:10.1006/jmla.2001.2798

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. doi:10.3758/Bf03193501

Hyde, T. S., & Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior, 12*(5), 471-480. doi:10.1016/S0022-5371(73)80027-1

Jacoby, L. L., & Bartz, W. H. (1972). Rehearsal and transfer to LTM. *Journal of Verbal Learning and Verbal Behavior, 11*, 561-656. doi:10.1016/S0022-5371(72)80039-2

Johnson, M. K. (1992). MEM: Mechanisms of Recollection. *Journal of Cognitive Neuroscience, 4*(3), 268-280. doi:10.1162/jocn.1992.4.3.268

Johnson, M. K., Reeder, J. A., Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychological Science, 13*(1), 64-67. doi:10.1111/1467-9280.00411

Levy, B. A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*(2), 123-132. doi:10.1016/S0022-5371(71)80003-8

Loaiza, V. M., Duperreault, K. A., Rhodes, M. G., & McCabe, D. P. (2015). Long-term semantic representations moderate the effect of attentional refreshing on episodic memory. *Psychonomic Bulletin & Review, 22*(1), 274-280. doi:10.3758/s13423-014-0673-7

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. doi:10.3758/s13421-011-0148-2

Loaiza, V. M., & McCabe, D. P. (2013). The influence of aging on attentional refreshing and articulatory rehearsal during working memory on later episodic memory performance. *Aging Neuropsychology and Cognition, 20*(4), 471-493. doi:10.1080/13825585.2012.738289

Loaiza, V. M., McCabe, D. P., Youngblood, J. L., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1258-1263. doi:10.1037/a0023923

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, Series B: Psychological Sciences and Social Sciences, 70(5), 729-736. doi:10.1093/geronb/gbt119

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science, 20*(4), 240-245. doi:10.1177/0963721411415340

Martin, R. C. (2005). Components of short-term memory and their relation to language processing - Evidence from neuropsychology and neuroimaging. *Current Directions in Psychological Science, 14*(4), 204-208. doi:10.1111/j.0963-7214.2005.00365.x

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1-23. doi:10.3758/s13428-011-0124-6

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*(3), 679-690. doi:10.3758/s13428-010-0049-5

Mazuryk, G. F., & Lockhart, R. S. (1974). Negative Recency and Levels of Processing in Free-Recall. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 28*(1), 114-123. doi:10.1037/h0081971

Melton, A. W. (1963). Implications of short-term-memory for a general-theory of memory. *Journal of Verbal Learning and Verbal Behavior, 2*(1), 1-21. doi:10.1016/S0022-5371(63)80063-8

Nairne, J. S. (1990). A Feature Model of Immediate Memory. *Memory & Cognition, 18*(3), 251-269. doi:10.3758/Bf03213879

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81. doi:10.1146/annurev.psych.53.100901.135131

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique, 101*, 447-462.

Nishiyama, R. (2013). Dissociative Contributions of Semantic and Lexical-Phonological Information to Immediate Recognition. *Journal of Experimental Psychology: Learning Memory and Cognition, 39*(2), 642-648. doi:10.1037/a0029160

Nishiyama, R. (2014). Active maintenance of semantic representations. *Psychonomic Bulletin & Review, 21*(6), 1583-1589. doi:10.3758/s13423-014-0618-1

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 411. doi:10.1037//0278-7393.28.3.411

Oberauer, K. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition, 37*(3), 346-357. doi:10.3758/Mc.37.3.346

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex, 43*(1), 135-145. doi:10.1016/S0010-9452(08)70451-9

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*(2), 447-453. doi:10.1006/nimg.2001.0983

Raye, C. L., Mitchell, K. J., Reeder, J. A., Greene, E. J., & Johnson, M. K. (2008). Refreshing one of several active representations: Behavioral and functional magnetic resonance imaging differences between young and older adults. *Journal of Cognitive Neuroscience, 20*(5), 852-862. doi:10.1162/jocn.2008.20508

Rose, N. S., Buchsbaum, 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. doi:10.3758/s13421-014-0398-x

Rose, N. S., & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning Memory and Cognition, 38*(4), 1019-1029. doi:10.1037/a0026976

Rose, N. S., Craik, F. I. M., & Buchsbaum, B. R. (2015). Levels of Processing in Working Memory: Differential Involvement of Frontotemporal Networks. *Journal of Cognitive Neuroscience, 27*(3), 522-532. doi:10.1162/jocn_a_00738

Rose, N. S., Myerson, J., Roediger, H. L., & Hale, S. (2010). Similarities and differences between working memory and long-term memory: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning Memory and Cognition, 36*(2), 471-483. doi:10.1037/a0018405

Shivde, G., & Anderson, M. C. (2011). On the existence of semantic working memory: Evidence for direct semantic maintenance. *Journal of Experimental Psychology: Learning Memory and Cognition, 37*(6), 1342-1370. doi:10.1037/a0024832

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. doi:10.1037/0033-295x.114.1.104

Vergauwe, E., Camos, V., & Barrouillet, P. (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. doi:10.1037/a0035779

Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval rate may describe many cognitive procedures. *Frontiers in Human Neuroscience, 8*, 126. doi:10.3389/fnhum.2014.00126

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychonic Bulletin & Review, 22*(4), 1001-1006. doi:10.3758/s13423-014-0755-6

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 104*(4), 442-452.

Table 1. Model assumptions and predictions regarding the influence of semantic processing in retrieval from working memory.

	Predictions on the efficacy of semantic (S) vs. phonological (P) cues			
	Exp. 1	Exp. 2	Exp. 3	Exp. 4
Models Assumptions and *General Predictions*		Refresh vs. Rehearse Strategies	Low vs. High Cognitive Load	Intentional vs. Incidental Encoding
Original Multi-component Model (Baddeley, 1986)				
Maintenance and recall in WM is primarily phonological in nature due to rehearsal in the phonological loop *Memoranda actively maintained in WM should show a phonological cue benefit; a semantic retrieval cue benefit should only be evident for retrieval from LTM*	Phonological > Semantic	Phonological > Semantic regardless of strategy	Phonological > Semantic regardless of CL	Phonological > Semantic for Intentional Encoding; Semantic > Phonological for Incidental Encoding
Time-Based Resource Sharing Model (Barrouillet & Camos, 2015)				
Maintenance and recall in WM is primarily accomplished via refreshing by reconstructing decaying memory traces in an episodic buffer *A semantic retrieval cue benefit should be evident when refreshing is instructed or available (low cognitive load) during active maintenance of memoranda in WM*	Semantic > Phonological	Semantic > Phonological for refreshing strategy; Phonological > Semantic for rehearsal strategy	Semantic > Phonological for low CL; smaller or null for high CL	Semantic > Phonological for low CL, especially during intentional encoding
Covert Retrieval Model (McCabe, 2008)				
Maintenance and recall in WM is primarily accomplished via refreshing as a covert retrieval from LTM to reactivate displaced traces back into WM *A semantic retrieval cue benefit should be evident when refreshing is instructed or required (high cognitive load) during active maintenance of memoranda in WM*	Semantic > Phonological	Semantic > Phonological for refreshing strategy; smaller or null for rehearsal strategy	Semantic > Phonological for high CL; smaller or null for low CL	Semantic > Phonological for high CL, especially during intentional encoding

Embedded Processes Model (Vergauwe & Cowan, 2015)				
Maintenance and recall in WM is primarily accomplished via refreshing as a rapid scanning of active representations in the central component of WM embedded within LTM *A semantic retrieval cue benefit should be evident during active maintenance of memoranda in WM, regardless of the availability of refreshing*	Semantic > Phonological	Semantic > Phonological for the refreshing strategy	Semantic > Phonological regardless of CL	Semantic > Phonological regardless of CL, especially during intentional encoding
Unitary models of memory (e.g., Crowder, 1982; Nairne, 2002)				
WM recall is driven by cue-based retrieval from LTM; there is no distinction between WM and LTM *A semantic retrieval cue benefit should be evident regardless of factors intended to manipulate active maintenance in WM*	Semantic > Phonological	Semantic > Phonological regardless of strategy	Semantic > Phonological regardless of CL	Semantic > Phonological regardless of CL or encoding

Note. Exp. = experiment; WM = working memory; LTM = long-term memory; CL = cognitive load.

Table 2. Descriptive statistics (means and standard deviations) of recall, cue use, processing task accuracy and response times (RTs) across experiments.

		Recall		Cue Use		Processing Task Accuracy		Processing Task RT (ms)	
Experiment	Condition	Semantic	Phonological	Semantic	Phonological	Semantic	Phonological	Semantic	Phonological
1		.58 (.17)	.55 (.20)	.34 (.16)	.33 (.15)	.95 (.05)	.96 (.06)	377 (91)	377 (107)
2	Rehearsal	.61 (.18)	.58 (.19)	.31 (.15)	.33 (.18)	.98 (.02)	.98 (.02)	411 (80)	409 (74)
	Refreshing	.60 (.17)	.47 (.14)	.33 (.14)	.39 (.15)	.99 (.01)	.99 (.02)	399 (94)	410 (95)
3	Low Load	.55 (.22)	.54 (.20)	.34 (.19)	.36 (.17)	.97 (.04)	.97 (.03)	411 (47)	414 (48)
	High Load	.35 (.18)	.40 (.20)	.52 (.22)	.46 (.21)	.91 (.06)	.92 (.06)	566 (46)	574 (46)
4	Low Load Incidental	-	-	-	-	.93 (.09)	.92 (.24)	407 (59)	402 (51)
	Low Load Intentional	-	-	-	-	.92 (.08)	.92 (.08)	428 (38)	423 (57)
	High Load Incidental	-	-	-	-	.82 (.08)	.88 (.08)	559 (33)	539 (26)
	High Load Intentional	-	-	-	-	.80 (.10)	.83 (.11)	563 (36)	555 (44)

Note. Standard deviations are in parentheses.

Figure 1. Proportion of cued recall accuracy as a function of cue type in Experiments 1 – 3. Error bars represent one standard error of the mean.

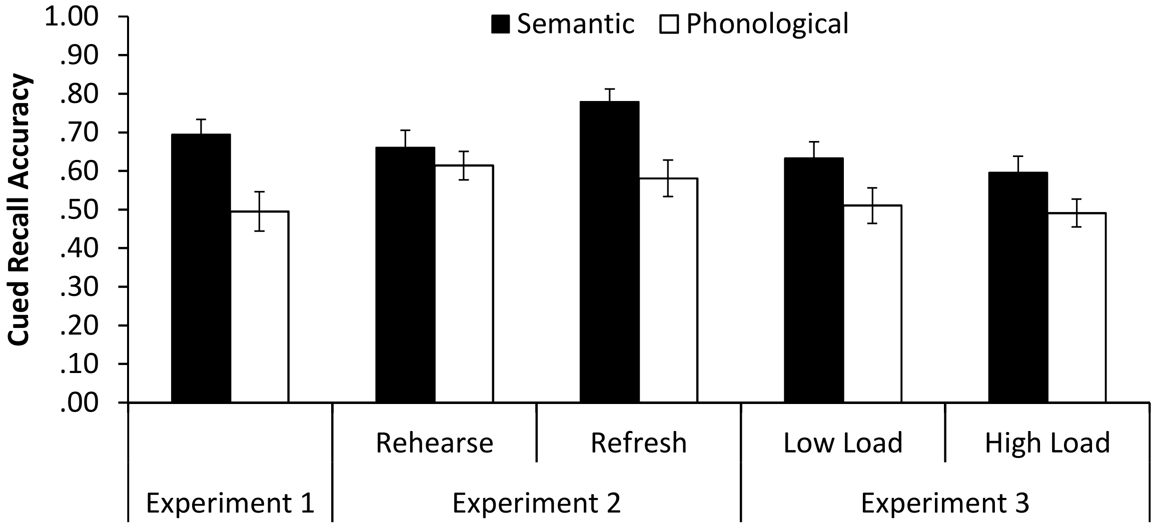

Figure 2. Proportion of cued recall accuracy as a function of encoding, cognitive load, and cue type in Experiment 4. Error bars represent one standard error of the mean.