Psychonomic Bulletin & Review (2024) 31:790–801
https://doi.org/10.3758/s13423-023-02381-4

BRIEF REPORT

How do intentions modulate the effect of working memory on long-term memory?

Maximilien Labaronne[1] · Laura Ferreri[1,2] · Gaën Plancher[1,3]

Accepted: 31 August 2023 / Published online: 18 September 2023

Abstract
Previous studies found that working memory maintenance contributes to long-term memory formation, and some evidence suggests that this effect could be larger when individuals are informed of the final long-term memory test. However, no study so far has explored whether and how working memory maintenance adapts when long-term retention is intentional. In this study, we conducted two experiments using verbal complex span tasks followed by delayed-recall tests. In both experiments, we evaluated working memory maintenance by varying the cognitive load of the concurrent task and with memory strategies reports. We manipulated intentions to remember at long term by warning participants of the final delayed recall or not (Experiment 1) or by monetarily rewarding immediate or delayed-recall performance (Experiment 2). We found no evidence that intentions changed the working memory maintenance mechanisms and strategies used, yet the cognitive load (Experiment 1) and rewards (Experiment 2) effects on delayed recalls were increased with a higher intention to remember at long term. We discuss possible interpretations for these results and suggest that the effect of intentions may not be due to a change in the kind of maintenance mechanisms used. As our results cannot be explained solely by encoding or maintenance processes, we instead propose that intentions produce a combined change in encoding and maintenance. However, the exact nature of this modulation will need further investigation. We conclude that understanding how intentions modulate the effect of working memory on long-term memory could shed new light on their relationship.

Keywords Working memory · Long-term memory · Intention · Reward

Introduction

Working memory maintenance and long-term retention

Researchers in the working memory (WM) field have shown increasing interest in the link between WM and long-term memory (LTM), in particular how one WM maintenance mechanism, attentional refreshing, contributes to LTM formation. Refreshing is defined as a domain-general maintenance mechanism that relies on attention to increase the activation level of information in WM and keep it accessible (Barrouillet & Camos, 2015; Camos et al., 2018; Johnson, 1992).

Attentional refreshing is mainly studied using complex span tasks. In these tasks, to-be-recalled items are presented sequentially and interspaced by a concurrent processing task (e.g., parity task, operation task). The Time-Based Resource-Sharing model (Barrouillet et al., 2004) proposes that refreshing availability is modulated by the cognitive load (CL) of the concurrent processing task, defined as the portion of the total time of the task during which attention is diverted from maintenance. CL can be manipulated by varying the number of distractors in the same period, the pace of the concurrent task, or its difficulty. Increasing CL reduces memory performance at immediate recall and, more critically, at delayed recall (Barrouillet et al., 2007; Camos & Portrat, 2015; Jarjat et al., 2018, 2020; Plancher & Barrouillet, 2013). Similarly, it has been observed that complex span tasks lead to better episodic memory than simple span tasks, supposedly by increasing the number of refreshing

✉ Maximilien Labaronne
m.labaronne@univ-lyon2.fr

1 Laboratoire d'Étude des Mécanismes Cognitifs, Université Lumière Lyon 2, Bron, France

2 Deparment of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy

3 Institut Universitaire de France (IUF), Paris, France

opportunities (the McCabe effect; Loaiza & McCabe, 2012; McCabe, 2008). Together, these results suggest that on top of short-term retention, refreshing also promotes LTM.

However, because the main purpose of WM is to allow short-term retention of information, it is not always relevant to remember this information at long term. Thus, we can wonder if intention to remember at long term modulates the effect of WM on LTM.

Role of recall relevance in working memory (WM) and long-term memory (LTM)

A meta-analysis on the effect of WM on LTM (Hartshorne & Makovski, 2019), highlighted that the effect of maintenance on long-term recall was stronger in studies that announced the upcoming delayed memory test, compared to studies using a surprise recall. This result suggests that intention to learn at long term increases the effect of WM on LTM. However, to our knowledge, the cause of this effect has not yet been investigated. One possibility could be that WM maintenance adapts to better promote LTM when aiming for long-term recall. Thus, we could expect that warning participants of a final delayed-recall test would lead to a change in the maintenance strategies deployed.

Another way to investigate intentions is to modulate participants' motivation by manipulating the relevance of information through external rewards, such as monetary or abstract values associated with the to-be-remembered information. There is substantial evidence regarding the effect of reward on WM. Financially rewarding correct recall in visual WM increases WM capacity for high- compared to low-reward trials (Kawasaki & Yamaguchi, 2013), and varying reward value at the item-level improves recall accuracy for high reward items (Allen & Ueno, 2018; Hitch et al., 2018). Similar results were found for verbal memory presented visually (Sandry et al., 2014, 2020) or auditorily (Atkinson et al., 2021). Reward boosts in WM have been proposed to be driven, at least partially, by the use of refreshing (Atkinson et al., 2022; Sandry et al., 2014, 2020; Sandry & Ricker, 2020).

Some studies also investigated whether WM reward-driven prioritization effects transfer to subsequent LTM. In Sandry et al. (2020) study, WM trials contained high-value items presented in red that were worth more points than low-value items presented in black. A final free delayed recall indicated better long-term recall for prioritized items, similar to what was observed in WM. Jeanneret et al. (2022) investigated the long-term consequences of two retrospective prioritization procedures in WM: reward-based (i.e., retrospectively assigned abstract value) and retro-cue (i.e., an arrow pointing to the previous location of the to-be-tested item). Using object images, they found that the benefit of cued items transferred to long-term recognition, while reward effects were more ambiguous. Together, these two studies suggest that prioritization in WM can also affects the resulting LTM of the information. However, and critically, these studies did not reward explicitly long-term recall but only the WM task, and used surprise LTM tests. It is not clear, therefore, whether the effects would be the same if participants would prioritize LTM formation during WM maintenance.

The present study

In the present study, we aimed to better understand the role of intentions on the link between WM and LTM by investigating (1) if the effect of WM on LTM is modulated by intention to remember at long term and (2) whether and how WM maintenance adapts to fulfill these intentions.

In a first experiment, we manipulated intention to remember at long term by informing or not participants of the final delayed-recall test before the WM task. We manipulated the CL of the concurrent task to evaluate the use of attentional refreshing (Barrouillet et al., 2007; Camos & Portrat, 2015; Jarjat et al., 2018, 2020; Plancher & Barrouillet, 2013). As it has been proposed that the addition of distractors can modulate the effect of WM on LTM (Loaiza & McCabe, 2012; McCabe, 2008), we also included a simple span condition to be compared with complex span conditions. If higher intention to remember at long term leads to increased use of attentional refreshing, we should observe a stronger CL effect at immediate and delayed recall when participants were informed of the delayed test, as compared to when they were not (i.e., lower intention to remember at long term).

In a second experiment, we modulated intentions using monetary rewards. We varied the value of the items to create a within-trial prioritization (i.e., high and low rewards associated with the to-be-encoded material, as in Hitch et al., 2018; Sandry et al., 2020), and the CL of the concurrent task as in Experiment 1 (excluding the simple span condition). Additionally, we manipulated which type of recall tests (i.e., immediate or delayed) yielded rewards to induce prioritization of either short- or long-term retention. Following previous literature, we anticipated that high-value items would be better recalled than low-value items at both immediate and delayed recall. In line with the hypotheses of Experiment 1, we expected that rewarding delayed recalls would lead to a stronger CL effect on immediate and delayed recalls than rewarding immediate recalls. We also predicted that the reward effect would be stronger on the rewarded test, immediate or delayed recall, compared to the non-rewarded test.

Experiment 1

Method

Participants

Eighty young adults (58 females) between 18 and 30 years old (M = 23.51, SD = 3.40 years) were recruited for this experiment. All participants had to be aged between 18 and 30 years old, be native French speakers, have normal or corrected-to-normal vision, and did not report any history of neurological or reading problems. All participants provided written informed consent before taking part in the study.

Material and design

The experiment was developed using jsPsych (de Leeuw, 2015) and was conducted online on a JATOS server (Lange et al., 2015). The experimental material consisted of 240 words that were selected from the Lexique3 database (New et al., 2001, 2004). Words were high frequency (M = 138.21, SD = 126.50, in written occurrences per million) singular common nouns, four to eight letters, and one to two syllables long. The concurrent parity task consisted of 32 sequences of 15 digits ranging from 1 to 9, with a pseudorandom presentation order.

At the end of the experiment, participants were presented a strategies form. This form asked participants to indicate the percentage of use during the WM task among the proposed strategies, the total having to reach 100%. Each strategy was described by a short sentence[1], and an additional fillable field ("other") could be used to report the use of non-proposed strategies. The proposed strategies were verbal rehearsal, attentional refreshing, stories, mental images, places (method of loci), mental line, and visual scenes. Finally, participants filled a motivation evaluation scale (Intrinsic Motivation Inventory; Ryan, 1982) composed of four dimensions: perceived choice, perceived competence, interest/enjoyment, pressure/tension.

We manipulated the maintenance condition of the task (simple span vs. complex span low CL vs. complex span high CL) as a within-subject variable and the delayed recall awareness (aware vs. unaware) as a between-subject variable. Two versions of the experiment were used to counterbalance words, digits, and CL condition order.

Procedure

The experiment started with the initial instructions describing the various phases of the experiment and the WM task. An additional sentence indicating the presence of a final delayed recall was added for half the participants (aware group) but not the other half (unaware group), and this recall was reminded in the halfway break screen for aware participants only. The experimental session then started with a training phase that was followed by the experimental task, and ended with the strategies form, the motivation evaluation scale and a final form asking for participants' personal information.

The experimental task was a simple or complex span task (Fig. 1). It consisted of 48 trials of five memoranda. Each trial started with a fixation cross for 1,000 ms. Words were presented sequentially for 1,500 ms each. In complex span trials, words were interspaced by a concurrent parity task consisting of three digits that participants had to judge as even or odd ("m" key for even and "q" for odd on an Azerty keyboard), that were presented for 600 ms (high CL) or 1,200 ms (low CL). In simple span trials, words were interspaced by a 3,000 ms black screen. Each word and digit was followed by an interstimulus interval of 150 ms. The time interval between two words in the simple span trials was selected to approximately match the available free time in the low CL condition. To estimate the time required to process a digit, we referred to previous literature that used a similar concurrent parity task within a complex span design. These studies reported presentation times ranging from 1,200 ms to 1,700 ms per digit, resulting in processing times of 621 ms and 615 ms in the two experiments conducted by Labaronne et al. (2023), and 672 ms and 625 ms for adults in the study by Rosselet-Jordan et al. (2022). Based on this information, we estimated the digit processing time to be approximately 650 ms. Considering this processing time, the free time between two words in the low CL condition was estimated to be 3,150 ms (subtracting three times 650 ms, to account for processing, to the total duration of 5,100 ms). Consequently, we used this duration for the simple span task (150 ms of interstimulus interval after the word, followed by a 3,000-ms blank screen). Participants were not asked to read the words or digits aloud.

At the end of a trial, a recall screen prompted participants to recall the five presented words in their original order by typing in the five response boxes, without time limit. The recall was confirmed by clicking on a submit button and was followed by a message asking participants to place again their fingers on the *m* and *q* keys to prepare for the coming trial. The next trial was started by pressing the spacebar.

[1] These were the sentences presented to the participants (translated from French). Verbal rehearsal: "Verbally rehearse in your head/orally". Mental images: "Form separate mental images for each word". Visual scene: "Imagine a visual scene that contains images of all words". Stories: "Telling yourself a story linking words". Mental line: "Place words or images on a mental line". Refreshing: "Think back to the words, but without saying them in your head". Places: "Mentally place the words in a familiar location". Other: "Other: please enter a description of the strategy".

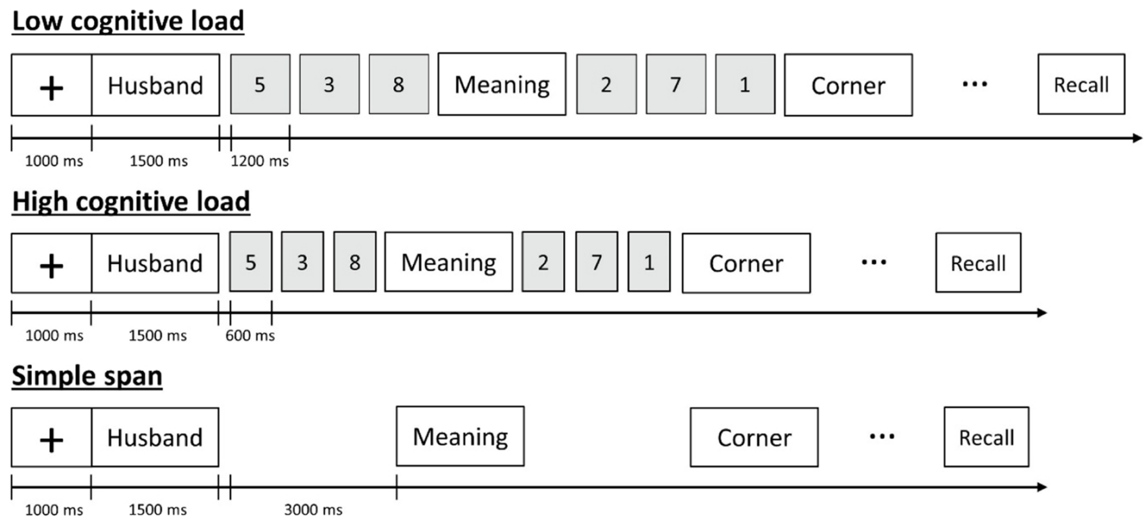

Fig. 1 Illustration of maintenance conditions in the task used in Experiment 1

Following the last trial, a 1-min distracting task of multiplication problems (e.g., 4 × 5 = 20?) was presented. After this distracting task, participants performed a delayed recall, in which they were invited to recall as many words as possible in any order and without time limit using the keyboard. This recall was confirmed by clicking on a submit button, followed by the motivation evaluation scale and the maintenance strategies form. A final form asked participants for their personal information, after what the experiment ended.

The practice was constituted of three phases. The first phase comprised 54 practice trials of the parity task without time limit. Accuracy was calculated, and the task had to be performed again if it did not reach 70%. The second phase corresponded to ten examples of the arithmetic problems used in the distracting task before the delayed-recall test. The third phase was similar to the experimental task, consisting of six trials so that every condition of CL was presented twice along with two simple span trials. Performance on the secondary parity task was monitored during the entire experiment. If their correct-response rate reached the lower limit of 70%, participants were reminded the importance of paying attention to this task and were warned that they had to increase their performance to prevent being sent back to the practice phase. A new performance check was done three trials later, and participants were indeed required to complete again the parity practice phase if their performance had not increased above 70%.

Statistical analyses

To ensure active processing of the concurrent parity task, participants having less than 70% of correct responses were excluded from analyses (final $n = 74$) as done in some previous studies (e.g., Camos et al., 2011, 2019; Labaronne et al., 2023).

Analyses were conducted using R (R Core Team, 2019) and RStudio (RStudio Team, 2021) with BayesFactor (Morey & Rouder, 2015) and bayestestR (Makowski et al., 2019) packages. Bayesian analyses of variance were conducted on immediate serial and delayed recall scorings, using maintenance condition (simple span vs. complex span low CL vs. complex span high CL) and delayed recall awareness (aware vs. unaware) as predictors and subjects as a random factor. Bayesian models were compared to a null model including only a random effect of subjects. The likelihood of each effect was assessed using $BF_{inclusion}$ and $BF_{exclusion}$ calculated across matched models (Mathôt, 2017), reflecting the proofs in favor or against an effect. Resulting BFs were interpreted using the following classification (Lee & Wagenmakers, 2013, adapted from Jeffreys, 1961): BF at 1 shows no evidence, anecdotal evidence between 1 and 3, moderate evidence between 3 and 10, strong evidence between 10 and 30, very strong evidence between 30 and 100 and extreme evidence for BF > 100. Therefore, BFs below 3 were interpreted as inconclusive. Interactions supported by at least moderate evidence were decomposed using Bayesian t-tests.

By design, strategies values were interdependent, as response to a strategy conditioned possible values to the other strategies, which made it unsuited for statistical analyses. Thus, we observed descriptively if the strategies data suggested any difference between aware and unaware participants.

Additional exploratory analyses that were conducted are included as Appendices. First, to evaluate if the effect of intentions could depend on individuals' WM capacity, the median of the simple span performance was used to distinguish high- and low-performance participants but it did

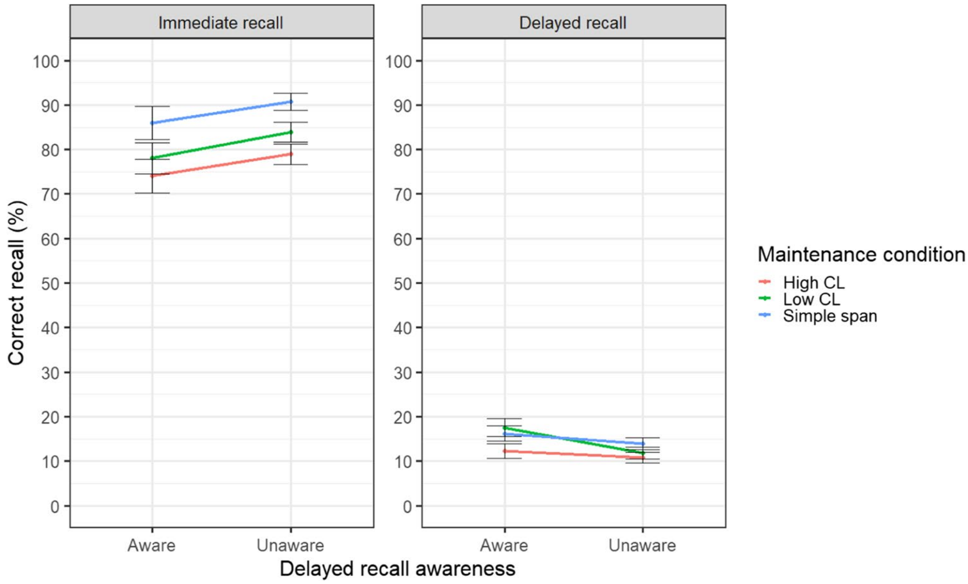

Fig. 2 Mean correct recall percentage in Experiment 1 by maintenance condition of the task (simple span vs. complex span low vs. complex span high) and delayed recall awareness group (aware vs. unaware). The error bars refer to the standard error

not affect the other factors (Appendix 1). To explore a possible role of motivation level, we tested if differences in the reported motivation could be found between the two groups or if the scores in some dimensions correlated with memory performance at short- or long-term. We mostly found evidence supporting the null or inconclusive results (Appendix 2).

Results and discussion

Main analysis

At immediate recall (Fig. 2), the likeliest model included only the main effect of maintenance condition (BF_{10} = 1.62e+15). There was extreme evidence in favor of the effect of maintenance condition ($BF_{inclusion}$ = 1.63e+15), with better recall in simple span trials (*M* = 88.26, *SD* = 18.36) than under low (*M* = 80.88, *SD* = 18.10) and high (*M* = 76.42, *SD* = 19.29) CL. Evidence for the main effect of awareness was inconclusive ($BF_{exclusion}$ = 1.44), and there was strong evidence against the maintenance condition × awareness interaction ($BF_{exclusion}$ = 10.63). Thus, we did not find evidence that awareness of the upcoming delayed-recall test affected overall immediate performance or interacted with maintenance condition at short-term. A complementary analysis conducted on a lenient immediate recall scoring, that did not consider items' serial position, led to similar findings (Appendix 3).

At delayed recall, the likeliest model included the main effects of maintenance condition and awareness, and the maintenance condition × awareness interaction (BF_{10} = 675.80). There was extreme evidence for the effect of maintenance condition ($BF_{inclusion}$ = 454.58), with better recall in simple span (*M* = 15.08, *SD* = 9.43) and low CL trials (*M* = 14.73, *SD* = 10.67) compared to high CL trials (*M* = 11.54, *SD* = 8.75). Evidence for the main effect of awareness ($BF_{exclusion}$ = 1.06) and the maintenance condition × awareness interaction ($BF_{inclusion}$ = 1.58) was inconclusive. Given it was included in the best model and its importance for our theoretical hypothesis, we conducted post hoc Bayesian t-tests on the maintenance condition × awareness interaction. In the unaware group, the difference between simple span and low CL was inconclusive (BF_{10} = 1.09). However, evidence supported a difference between simple span and high CL (BF_{10} = 3.12) and an absence of difference between low and high CL (BF_{01} = 3.47). In the aware group, evidence pointed against a difference between simple span and low CL (BF_{01} = 3.26) but supported a difference between simple span and high CL (BF_{10} = 10.65) and between low and high CL (BF_{10} = 362.66). We did not replicate the classical McCabe effect (Loaiza & McCabe, 2012; McCabe, 2008) by which items studied in complex span trials are better recalled at long term than those studied in simple span trials. This difference may be caused by the addition of free time between memoranda in simple span trials (e.g., Souza & Oberauer, 2017). As attention is not displaced by distractors in simple span, it may also be that this free time was used for short-term consolidation, which has been shown to benefit both WM and LTM and is independent from maintenance mechanisms (Cotton & Ricker, 2021; Labaronne et al., 2023). It is thus difficult to clearly conclude on the simple span results. Interestingly, we replicated the delayed

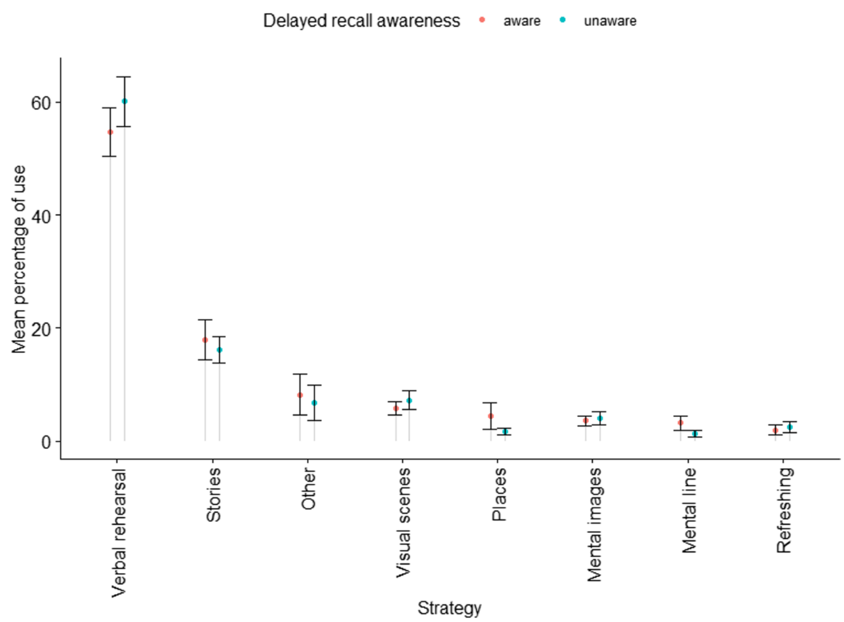

Fig. 3 Mean percentage of strategy use in Experiment 1 depending on delayed recall awareness. Error bars refer to the standard error

CL effect in complex span tasks (Camos & Portrat, 2015; Jarjat et al., 2018, 2020) in the aware group (+5.23), but not in the unaware group (+0.88). We should note that we did not control for whether participants predicted the LTM test. This could explain the inconclusive evidence regarding awareness and the interaction between awareness and maintenance condition, however results of t-tests on the CL effect at delayed recall goes against this hypothesis. Future studies should examine this possibility more carefully.

Strategies

Descriptive observation of the data did not suggest a clearly distinct pattern of strategies between the two groups, at best only minor differences (Fig. 3).

Results summary

Surprisingly, the results of this experiment suggest that the knowledge of an upcoming delayed recall does not visibly affect WM maintenance, neither in recall performance nor in maintenance strategies deployed, but still modulates the impact of complex span task CL on LTM.

Experiment 2

A second experiment was conducted to test our theoretical hypothesis using different manipulations. Because it has been shown that rewards in WM also affect subsequent LTM (Sandry et al., 2020), we investigated if the long-term reward effect could be modulated by reward-driven intentions regarding long-term retention. We manipulated monetary rewards within-trials using high- and low-value memoranda as in previous literature (e.g., Hitch et al., 2018; Sandry et al., 2020), and intentions were manipulated by varying whether rewards were given for performance in immediate or delayed memory tests.

Method

Participants

As participants were financially compensated, the sample size was reduced compared to Experiment 1 due to resources limitations. We recruited 60 new participants (43 females) aged between 18 and 28 years old (M = 21.92, SD = 2.23) for this experiment. The recruitment criteria were the same as in Experiment 1. Although participants believed that their reward would depend on their memory performance during the task, all those who completed the experiment received 10€ for their participation.

Material and design

We randomly selected 140 items from Experiment 1 material. Given the absence of effect of intentions on motivation levels when manipulating delayed recall awareness in Experiment 1 (Appendix 2) or rewards in previous literature (Sandry & Ricker, 2020), the motivation evaluation scale was removed for this experiment. Additionally, to prevent the response to one strategy from being dependent on the responses to the other strategies, so we could use statistical analyses, the strategies form now asked for a percentage of use for each proposed strategy independently.

Procedure

The same procedure as in Experiment 1 was used, excepted for the following points. The experiment consisted of two blocks of 14 trials, with a delayed recall at the end of each block. At the beginning of the experiment, all participants were informed of the presence of delayed-recall tests. Participants were split into two groups, who were told that their reward (i.e., the final amount of money won) depended on their memory performance in either immediate or delayed-recall tests. The second or third memoranda in each trial, depending on the block, was a high-value item that rewarded 0.16€ upon correct recall while the rest of the items were rewarded 0.05€. The high-value item was presented underlined to ensure correct identification[2]. The high-value item's serial position in the first block was counterbalanced between participants (2 or 3) and changed between blocks. These instructions were reminded at the end of the training session and after the first block. After a rewarded test (immediate or delayed recall, depending on the group), a screen displayed the amount rewarded for this recall and the total gain. Given the conclusions of Experiment 1, the simple span condition was removed.

In summary, we manipulated the CL of the concurrent parity task (low vs. high) and the item reward value (high vs. low) as within-subject variables, and rewarded test (immediate recall vs. delayed recall) as a between-subject variable.

Statistical analyses

Analyses were conducted similarly to Experiment 1. Participants having less than 70% correct responses to the concurrent parity task were discarded from following analyses (final $n = 52$).

Bayesian analyses of variance (bANOVA) were conducted on mean correct recall percentage at both immediate and delayed recall, using CL (low vs. high), reward value (low vs. high) and rewarded test (immediate recall vs. delayed recall) as predictive variables, and subject as a random factor. Regarding the reward value factor, we computed the mean percentage of high-value items correctly recalled and the percentage of low-value correctly recalled, independently of their position in the trials.

[2] While it could be argued that underlying high-value items made them more perceptually distinct, previous studies have suggested that distinctiveness could not account for the value effect (e.g., Atkinson et al., 2021; Sandry et al., 2020; Sandry & Ricker, 2020).

For strategies data, the change made in this experiment removed the responses' interdependence and allowed us to use statistical analyses. We conducted a Bayesian analysis of variance on the percentage of use of strategies, with the rewarded test and the strategies as predictors, and subject as a random variable.

Results and discussion

Main analysis

At immediate recall (Fig. 4), the likeliest model included the main effects of CL and reward value (BF_{10} = 2.98e+09). Extreme evidence supported an effect of CL ($BF_{inclusion}$ = 8.42e+04), with better recall under low ($M = 88.85$, $SD = 12.26$) than high ($M = 83.21$, $SD = 14.14$) CL. Additionally, we found extreme evidence for a main effect of reward value ($BF_{inclusion}$ = 3.31e+05), with a better recall for high-value items ($M = 90.66$, $SD = 9.17$) than for low-value items ($M = 84.87$, $SD = 13.97$), replicating the reward effect in verbal WM (Atkinson et al., 2021; Sandry et al., 2014, 2020). Evidence regarding the main effect of rewarded test was inconclusive ($BF_{exclusion} = 2.81$). There was moderate evidence against the CL × reward value interaction ($BF_{exclusion} = 4.93$). As the effect of rewards was not modulated by the CL, thought to determine refreshing availability, this does not support the previous proposition (Atkinson et al., 2022; Sandry et al., 2014, 2020; Sandry & Ricker, 2020) that the effect of reward can be attributed to refreshing. CL did not interact with rewarded test either ($BF_{exclusion} = 3.45$), suggesting, consistent with Experiment 1, that intentions did not modulate the use of attentional refreshing. There was moderate evidence against the rewarded test × reward value interaction ($BF_{exclusion} = 3.28$) and the three-way interaction ($BF_{exclusion} = 3.38$). The analysis using lenient scoring led to similar findings (Appendix 3).

At delayed recall, the best model included the main effects of CL, reward value and rewarded test, and the rewarded test × reward value interaction (BF_{10} = 2.54e+12). There was extreme evidence for an effect of CL ($BF_{inclusion}$ = 5.08e+04), with better recall under low ($M = 22.42$, $SD = 14.24$) than high ($M = 16.37$, $SD = 10.91$) CL. Moderate evidence supported the main effect of rewarded test ($BF_{inclusion} = 3.15$), with better recall performance when the delayed recall was rewarded ($M = 22.50$, $SD = 12.73$) than when immediate recall was rewarded ($M = 16.29$, $SD = 10.38$). Consistent with the study of Sandry et al. (2020), results indicated extreme evidence for the main effect of reward value ($BF_{inclusion}$ = 1.00e+07). As with immediate recall scoring, there was moderate evidence against the CL × reward value ($BF_{exclusion} = 4.65$) and CL × rewarded test ($BF_{exclusion} = 3.66$) interactions. Crucially, the effect of reward value was stronger in the group in which

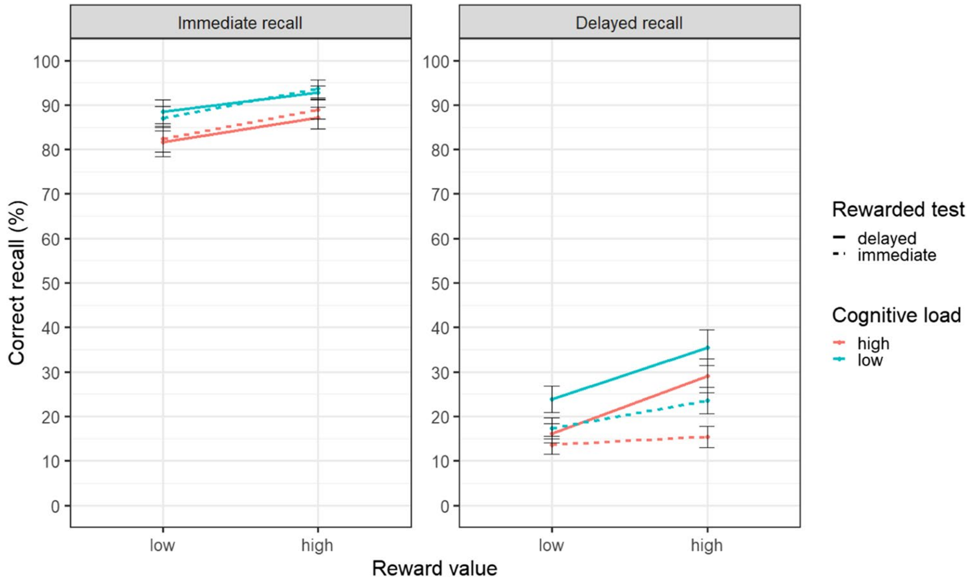

Fig. 4 Mean correct recall percentage in Experiment 2 by cognitive load of the concurrent task (low vs. high), item reward value (low vs. high) and rewarded test (immediate recall vs. delayed recall). The error bars refer to the standard error

delayed recall was rewarded (+12.23) compared to the group in which immediate recall was rewarded (+4.02), as suggested by the very strong evidence in favor of the reward value × rewarded test interaction[3] ($BF_{inclusion}$ = 36.97). Post hoc comparisons indicated inconclusive evidence regarding the effect of rewards in the immediate rewards group (BF_{01} = 1.82), and moderate evidence in the delayed rewards group (BF_{10} = 5.64). Evidence regarding the three-way interaction was inconclusive ($BF_{exclusion}$ = 1.97).

Strategies analysis

The best model included only the main effect of strategies (BF_{10} = 2.95e+25). We found moderate evidence against an effect of the rewarded test ($BF_{exclusion}$ = 6.76) and very strong evidence against the strategy × rewarded test interaction ($BF_{exclusion}$ = 66.67). As can be seen in Fig. 5, these results do not support the use of distinct maintenance strategies between the two groups.

Results summary

We found no evidence that manipulating the rewarded test (i.e., immediate or delayed recall) affected short-term recall. However, participants that were rewarded in the delayed-recall tests showed better recall performance and a stronger reward effect in delayed recalls. This experiment suggests that higher intention to remember at long term increases the long-term effect of rewards without modulating short-term recall.

General discussion

Our first objective was to investigate if intentions modulate the effect of WM on LTM. In Experiment 1, in line with the literature (Hartshorne & Makovski, 2019), we observed that awareness of the upcoming delayed recall increased the CL effect on delayed-recall performance, however the bANOVA gave inconclusive evidence and this effect was supported only by t-tests. In Experiment 2, we found that the long-term benefit of rewards was increased when rewarding delayed recall rather than immediate recall. Together, these first results suggest that intentions regarding long-term remembering could modulate the long-term recall of items presented in a WM task. Future studies will be needed to explore this question in more detail.

Our second objective was to observe if and how WM maintenance adapts when aiming for long-term recall. While we replicated the effect of CL on immediate recall (Barrouillet et al., 2007; Camos & Portrat, 2015; Jarjat et al., 2018, 2020; Plancher & Barrouillet, 2013) in both experiments, we failed to find evidence that this effect was modulated by intentions manipulation, would it be through delayed recall awareness or the rewarded test. Additionally,

[3] To evaluate if participants rewarded based on LTM performance recalled more words overall or only more high-value items, we also compared the two groups for low-value and high-value items distinctively. Evidence was inconclusive for low-value items (BF_{01} = 1.50), but moderate evidence supported a difference for high-value items (BF_{10} = 7.39). Thus, the rewarded test's effect seems to mainly come from a difference in recalling high-value items.

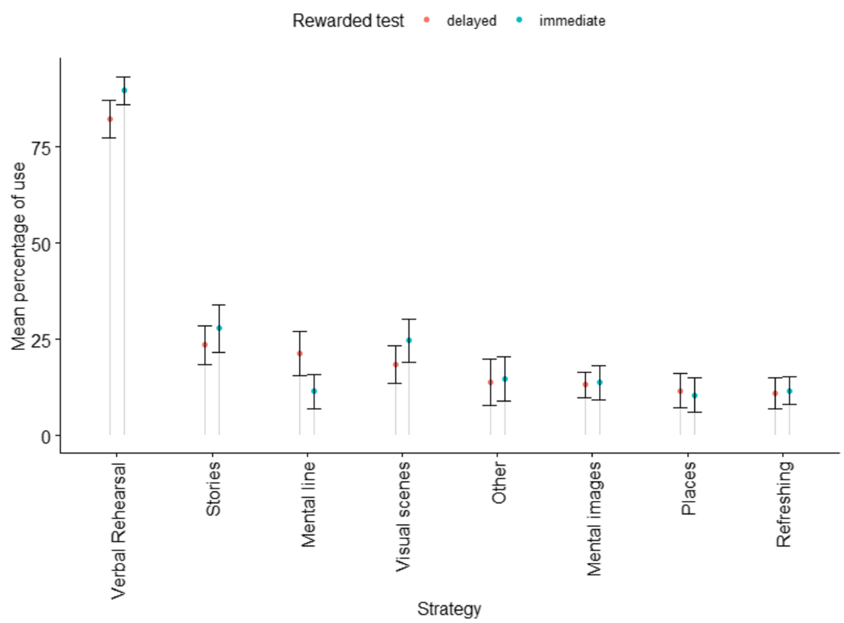

Fig. 5 Mean percentage of strategy use in Experiment 2 depending on the rewarded test. Error bars refer to the standard error

in both experiments, we did not find evidence that the maintenance strategies used differed. In sum, our results do not suggest that intention to remember at long term modulated the maintenance mechanisms used or their efficiency at short-term.

While our study provides new evidence that WM manipulations modulate subsequent LTM more strongly with higher intention to remember at long term, the cause of this effect remains to be identified. We anticipated that trying to learn at long term could lead to increased use of refreshing. The increased CL effect with awareness in the delayed recall of Experiment 1 could partly support this idea. However, refreshing is more sensitive to concurrent attentional demand than other low-attentional mechanisms such as verbal rehearsal (Camos, 2015; Camos et al., 2011; Camos & Barrouillet, 2014; Mora & Camos, 2013). Thus, intentions not modulating the effect of CL on immediate recall in both experiments makes this interpretation implausible. Alternatively, conditional use of elaborative strategies could be proposed to explain our results, as previous literature suggested that elaboration improves delayed recall without impacting immediate recall (Bartsch et al., 2018). However, this explanation requires to consider that their use is concurrent to other maintenance mechanisms affecting short-term recall to explain the short-term CL effect and, as we saw no change in the reported strategies, that participants were unable to explicitly report the use of such strategies. Similarly, it has been reported that a change in strategies was not able to explain the reward effect (Sandry & Ricker, 2020). Thus, the proposition of a change in the maintenance strategies used does not seem promising.

Considering our results and the literature, we propose the hypothesis that intentions induce a combined change in encoding and maintenance. In a long-term recognition task, a concurrent random number generation during encoding, contrary to an articulatory suppression task, removes the reward effect on delayed recognition (Elliott & Brewer, 2019). Coherently, imaging studies reported a preparatory phase before the presentation of items that can predict their successful encoding (Adcock et al., 2006; Addante et al., 2011, 2015), and that intention to remember modulates frontal-midline's low beta band power during this phase (Schneider & Rose, 2016). Thus, it appears that intention to remember at long term could alter encoding, leading to a better long-term recall. However, evidence also suggest a role of WM maintenance. In Experiment 1, we found that CL, a manipulation constricted to WM maintenance, affected delayed recall differently depending on delayed recall awareness. In Experiment 2, in which all participants were aware of the delayed test, CL affected the delayed recall of all participants. Accordingly, oscillatory activity during maintenance can predict both immediate and delayed-recall performance (Khader et al., 2010). Therefore, one possible explanation could be that intention to remember at long term does not modulate the kind and amount of maintenance mechanisms engaged. Instead, it could lead to a change in encoding processes, for instance through the nature of the encoded representation, that would determine to what extent WM maintenance mechanisms affect long-term recall. Further studies are needed to clarify how intentions modulate the effect of WM on LTM, which could bring a better understanding of the links between WM and LTM.

Appendices

Appendix 1

Simple span performance split in Experiment 1

First, we used the median of the simple span performance to distinguish high- and low-performance participants. In particular, we were interested to see if the interaction between CL and delayed recall awareness on delayed recall was modulated by participants' overall WM performance. Participants having a mean correct response in the simple span equal or inferior to the median of the means correct percentage were labeled as low performance, while the rest of the participants were labeled high performance. Simple span trials were then removed from the data set, and analyses were conducted similarly to the main analyses excepted for the inclusion of performance group as a factor. Two Bayesian analyses of variance were conducted on immediate serial recall and delayed recall, using CL (low vs. high), delayed recall awareness (aware vs. unaware), and performance group (low vs. high) as predictors and subjects as a random factor. Bayesian models were compared to a null model including only a random effect of subjects.

At immediate recall, the best model included the main effects of delayed recall awareness, CL, and performance group (BF_{10} = 2.62e+05). Evidence regarding the interaction between delayed recall awareness and performance group ($BF_{exclusion}$ = 2.58) and the interaction between the three factors ($BF_{exclusion}$ = 2.87) were inconclusive. There was moderate evidence supporting an absence of interaction between CL and performance group ($BF_{exclusion}$ = 4.02).

At delayed recall, the best model included the main effects of delayed recall awareness, CL, and performance group, the interaction between delayed recall awareness and CL, and the interaction between delayed recall awareness and performance group (BF_{10} = 5.08e+04). Evidence for the interaction between delayed recall awareness and performance group ($BF_{inclusion}$ = 1.10) and for the interaction between the three factors ($BF_{exclusion}$ = 2.73) were inconclusive. There was moderate evidence supporting the absence of interaction between CL and performance group ($BF_{exclusion}$ = 3.32).

Appendix 2

Motivation scale results in Experiment 1

First, we assessed if the reported scores in each of the four dimensions of the scale (perceived choice, perceived competence, interest/enjoyment, pressure/tension) differed between aware and unaware participants. We conducted Bayesian t-tests on each score between the two groups. Evidence supported an absence of difference for the perceived choice (BF_{01} = 3.25), perceived competence (BF_{01} = 3.86), and pressure (BF_{01} = 4.28) dimensions. The difference on the interest dimension was inconclusive (BF_{01} = 2.13).

Additionally, we evaluated if memory performance in immediate or delayed recalls correlated with dimensions on the motivation scale. One Bayesian correlation was conducted on each combination of motivation dimension and mean percentage of recall for immediate serial and delayed recall. Moderate evidence supported an absence of correlation between perceived choice and immediate serial recall performance (BF_{01} = 3.73). Evidence for a correlation between immediate serial recall and perceived competence (BF_{10} = 1.02), interest (BF_{01} = 2.52), and pressure (BF_{10} = 1.29) were inconclusive. There was moderate evidence supporting a correlation between delayed-recall performance and interest (BF_{10} = 3.14, r = .33). Evidence for a correlation between delayed recall and perceived choice (BF_{10} = 1.80), perceived competence (BF_{10} = 1.40), and pressure (BF_{10} = 1.99) were inconclusive.

Appendix 3

Lenient immediate recall analyses

We computed a lenient immediate recall scoring, which is similar to serial immediate scoring but does not take into account the serial position of the recalled items. Thus, an item correctly recalled but at a wrong serial position was considered incorrect in the serial scoring but correct in the lenient scoring. The analysis on the lenient scoring was conducted similarly to the analysis on serial immediate scoring in both experiments and led to similar findings.

In Experiment 1, the best model using the lenient scoring included only the main effect of maintenance condition (BF_{10} = 5.42e+12). Evidence for the main effect of delayed-recall awareness was inconclusive ($BF_{exclusion}$ = 1.51). There was extreme evidence supporting a main effect of maintenance condition ($BF_{inclusion}$ = 5.02e+12). There was strong evidence supporting an absence of interaction between maintenance condition and delayed-recall awareness ($BF_{exclusion}$ = 11.76).

In Experiment 2, the best model included the main effects of CL and reward value (BF_{10} = 7.09e+09). There was extreme evidence for an effect of CL ($BF_{inclusion}$ = 2.67e+03). There was moderate evidence for an absence of interaction between CL and reward value ($BF_{exclusion}$ = 4.50). CL also did not interact with rewarded test ($BF_{exclusion}$ = 3.58). Evidence regarding the interaction between the three factors was inconclusive ($BF_{exclusion}$ = 2.95). Additionally, we found extreme evidence for a main effect of reward value

($BF_{inclusion}$ = 2.16e+07). Results indicated inconclusive evidence for the main effect of rewarded test ($BF_{exclusion}$ = 2.35). Evidence against the interaction between rewarded test and reward value was moderate ($BF_{exclusion}$ = 4.17).

Funding This work was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). This research was supported in part by a grant from the Agence Nationale de la Recherche (ANR-18-CE28-0012) awarded to Gaën Plancher, Project REFLECTOR and by the Institut Universitaire de France.

Data availability Experiment files, datasets, and analyses scripts for the experiments are available via the Open Science Framework at: https://doi.org/10.17605/OSF.IO/Y4672.

Declarations

Competing interests The authors have no competing interest to declare.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron, 50*(3), 507–517. https://doi.org/10.1016/j.neuron.2006.03.036

Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., & Ranganath, C. (2011). Prestimulus theta activity predicts correct source memory retrieval. *Proceedings of the National Academy of Sciences, 108*(26), 10702–10707. https://doi.org/10.1073/pnas.1014528108

Addante, R. J., de Chastelaine, M., & Rugg, M. D. (2015). Pre-stimulus neural activity predicts successful encoding of inter-item associations. *NeuroImage, 105*, 21–31. https://doi.org/10.1016/j.neuroimage.2014.10.046

Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention, Perception, & Psychophysics, 80*(7), 1731–1743. https://doi.org/10.3758/s13414-018-1543-6

Atkinson, A. L., Allen, R. J., Baddeley, A. D., Hitch, G. J., & Waterman, A. H. (2021). Can valuable information be prioritized in verbal working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(5), 747–764. https://doi.org/10.1037/xlm0000979

Atkinson, A. L., Oberauer, K., Allen, R. J., & Souza, A. S. (2022). Why does the probe value effect emerge in working memory? Examining the biased attentional refreshing account. *Psychonomic Bulletin & Review, 29*(3), 891–900. https://doi.org/10.3758/s13423-022-02056-6

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction* (1st ed.). Psychology Press. https://doi.org/10.4324/9781315755854

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. https://doi.org/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. https://doi.org/10.1037/0278-7393.33.3.570

Bartsch, L. M., Singmann, H., & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory & Cognition, 46*(5), 796–808. https://doi.org/10.3758/s13421-018-0805-9

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

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: The executive and phonological loops. *Frontiers in Human Neuroscience, 8*, 900. https://doi.org/10.3389/fnhum.2014.00900

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

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

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory?: What is attentional refreshing? *Annals of the New York Academy of Sciences, 1424*(1), 19–32. https://doi.org/10.1111/nyas.13616

Camos, V., Mora, G., Oftinger, A.-L., Mariz Elsig, S., Schneider, P., & Vergauwe, E. (2019). Does semantic long-term memory impact refreshing in verbal working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000657

Cotton, K., & Ricker, T. J. (2021). Working memory consolidation improves long-term memory recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(2), 208–219. https://doi.org/10.1037/xlm0000954

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Elliott, B. L., & Brewer, G. A. (2019). Divided attention selectively impairs value-directed encoding. *Collabra: Psychology, 5*(1), 4. https://doi.org/10.1525/collabra.156

Hartshorne, J. K., & Makovski, T. (2019). The effect of working memory maintenance on long-term memory. *Memory & Cognition, 47*(4), 749–763. https://doi.org/10.3758/s13421-019-00908-6

Hitch, G. J., Hu, Y., Allen, R. J., & Baddeley, A. D. (2018). Competition for the focus of attention in visual working memory: Perceptual recency versus executive control: Focus of attention in visual working memory. *Annals of the New York Academy of Sciences, 1424*(1), 64–75. https://doi.org/10.1111/nyas.13631

Jarjat, G., Hoareau, V., Plancher, G., Hot, P., Lemaire, B., & Portrat, S. (2018). What makes working memory traces stable over time?: Working memory conditions for long-term recall. *Annals of the New York Academy of Sciences, 1424*(1), 149–160. https://doi.org/10.1111/nyas.13668

Jarjat, G., Plancher, G., & Portrat, S. (2020). Core mechanisms underlying the long-term stability of working memory traces still work in aging. *L'Année psychologique, 120*(2), 203. https://doi.org/10.3917/anpsy1.202.0203

Jeanneret, S., Bartsch, L. M., & Vergauwe, E. (2022). *To be or not to be relevant: Comparing short- and long-term consequences across working memory prioritization procedures* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/vp953

Jeffreys, H. (1961). *Theory of probability*. Clarendon.

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

Kawasaki, M., & Yamaguchi, Y. (2013). Frontal theta and beta synchronizations for monetary reward increase visual working memory capacity. *Social Cognitive and Affective Neuroscience, 8*(5), 523–530. https://doi.org/10.1093/scan/nss027

Khader, P. H., Jost, K., Ranganath, C., & Rösler, F. (2010). Theta and alpha oscillations during working-memory maintenance predict successful long-term memory encoding. *Neuroscience Letters, 468*(3), 339–343. https://doi.org/10.1016/j.neulet.2009.11.028

Labaronne, M., Jarjat, G., & Plancher, G. (2023). Attentional refreshing in the absence of long-term memory content: Role of short-term and long-term consolidation. *Journal of Cognition, 6*(1), 5. https://doi.org/10.5334/joc.246

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies"" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLOS ONE, 10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

Lee, M. D., & Wagenmakers, E.-J. (2013). Bayesian cognitive modeling: A practical course. *Cambridge University Press*. https://doi.org/10.1017/CBO9781139087759

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

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software, 4*(40), 1541. https://doi.org/10.21105/joss.01541

Mathôt, S. (2017). Bayes like a Baws: Interpreting Bayesian repeated measures in JASP [Blog post]. *COGSCI*. https://www.cogsci.nl/blog/interpreting-Bayesian-repeated-measures-in-jasp. Accessed 23 May 2019.

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language, 58*(2), 480–494. https://doi.org/10.1016/j.jml.2007.04.004

Mora, G., & Camos, V. (2013). Two systems of maintenance in verbal working memory: Evidence from the word length effect. *PLoS ONE, 8*(7), e70026. https://doi.org/10.1371/journal.pone.0070026

Morey, R. D., & Rouder, J. N. (2015). BayesFactor (version 0.9. 10-2). *Computer Software*.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary french : LEXIQUE™. *L'année Psychologique, 101*(3), 447–462. https://doi.org/10.3406/psy.2001.1341

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers, 36*(3), 516–524. https://doi.org/10.3758/BF03195598

Plancher, G., & Barrouillet, P. (2013). Forgetting from working memory: Does novelty encoding matter? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(1), 110–125. https://doi.org/10.1037/a0028475

R Core Team. (2019). *R: A language and environment for statistical computing* (3.5.3). R Foundation for Statistical Computing. https://www.R-project.org/

Rosselet-Jordan, F. L., Abadie, M., Mariz-Elsig, S., & Camos, V. (2022). Role of attention in the associative relatedness effect in verbal working memory: Behavioral and chronometric perspectives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(11), 1571–1589. https://doi.org/10.1037/xlm0001102

RStudio Team. (2021). *RStudio: Integrated development for R*. RStudio Inc. http://www.rstudio.com/

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*(3), 450–461. https://doi.org/10.1037/0022-3514.43.3.450

Sandry, J., & Ricker, T. J. (2020). Prioritization within visual working memory reflects a flexible focus of attention. *Attention, Perception, & Psychophysics, 82*(6), 2985–3004. https://doi.org/10.3758/s13414-020-02049-4

Sandry, J., Schwark, J. D., & MacDonald, J. (2014). Flexibility within working memory and the focus of attention for sequential verbal information does not depend on active maintenance. *Memory & Cognition, 42*(7), 1130–1142. https://doi.org/10.3758/s13421-014-0422-1

Sandry, J., Zuppichini, M. D., & Ricker, T. J. (2020). Attentional flexibility and prioritization improves long-term memory. *Acta Psychologica, 208*, 103104. https://doi.org/10.1016/j.actpsy.2020.103104

Schneider, S. L., & Rose, M. (2016). Intention to encode boosts memory-related pre-stimulus EEG beta power. *NeuroImage, 125*, 978–987. https://doi.org/10.1016/j.neuroimage.2015.11.024

Souza, A. S., & Oberauer, K. (2017). Time to process information in working memory improves episodic memory. *Journal of Memory and Language, 96*, 155–167. https://doi.org/10.1016/j.jml.2017.07.002