Free Time, Sharper Mind: A Computational Dive into Working Memory Improvement

Benjamin Kowialiewski[1,2], Steve Majerus[1]

[1]Psychology & Neuroscience of Cognition Research Unit (PsyNCog), University of Liège, Belgium

[2]Fund for Scientific Research F.R.S.-FNRS, Brussels, Belgium

Correspondence concerning this article should be addressed to Benjamin Kowialiewski, Psychology & Neuroscience of Cognition Research Unit, 4000 Liège, Belgium. E-mail: bkowialiewski@uliege.be

Open Science statement:

All the data and codes have been made available on the Open Science Framework: https://osf.io/bc3qp/

Abstract

Extra free time improves working memory (WM) performance. This free-time benefit becomes larger across successive serial positions, a phenomenon recently labeled the “fanning-out effect”. Different mechanisms can account for this phenomenon. In this study, we implemented these mechanisms computationally and tested them experimentally. We ran three experiments that varied the time people were allowed to encode items, as well as the order in which they recalled them. Experiment 1 manipulated the free-time benefit in a paradigm in which people recalled items either in forward or backward order. Experiment 2 used the same forward-backward recall paradigm coupled with a distractor task at the end of encoding. Experiment 3 used a cued recall paradigm in which items were tested in random order. In all three experiments, the best-fitting model of the free-time benefit included (1) a consolidation mechanism whereby a just-encoded item continues to be re-encoded as a function of the total free-time available and (2) a stabilization mechanism whereby items become more resistant to output interference with extra free time. Mechanisms such as decay and refreshing, as well as models based on the replenishment of encoding-resources, were not supported by our data.

Keywords: working memory; free-time benefit; computational modeling

Introduction

In working memory (WM) tasks, people remember more information when they have more time to study them (e.g., Oberauer, 2022; Penney, 1975; Tan & Ward, 2008). It has been shown that the process of encoding items in WM takes some time (e.g., ~500ms for simple stimuli such as tones and letters, see for instance Jolicœur and Dell'Acqua, 1998). Our question concerns the post-encoding processes participants engage in when they are given additional time to process the items. We refer to this additional time as "free time", as participants are free to engage these additional processes during this time. This question is of fundamental importance for the theoretical modeling of WM as it pertains to the question of how a WM representation is maintained. The current study provides a comprehensive examination of several candidate processes that have been proposed for explaining the free-time benefit in WM, by deriving precise predictions for each process based on a computational modeling approach, and by testing the predictions against empirical data.

Current Explanations of the Free-Time Benefit

Different mechanisms, illustrated in **Figure 2**, have been proposed to explain the free-time benefit of WM. A first conceivable mechanism is encoding-resource (Popov & Reder, 2020). Each trial begins with a certain amount of resource. Each encoding step depletes the currently available resource by a constant proportion, and it is this proportion which determine encoding strength. Hence, encoding is proportional to the available resource: more resource means stronger encoding. After encoding, the resource gradually replenishes over time until the presentation of the next item, which in turn consumes the same proportion of the remaining resource, and so forth. This mechanism predicts a free-time benefit, as the extra free-time provided between memoranda allows a stronger replenishment of the resource, yielding to stronger encoding for the subsequent to-be-encoded items. One characteristic of this mechanism is that it acts on items proactively: when the resource replenishes, this improves encoding strength for the subsequent, but not the preceding, items.

A second mechanism is consolidation (Ricker et al., 2018; Ricker & Vergauwe, 2022). In this, the just-encoded item continues to be re-encoded for a longer duration, increasing its strength. In typical WM models, encoding is done by binding items to a positional context (see also modeling part). Therefore, during consolidation, this item-context binding is reinforced further for the just-encoded item.

A third mechanism which has never been explored is stabilization. After encoding an item, the additional free time is used to stabilize the just-encoded item, making it less sensitive to interference. One characteristic of the stabilization mechanism is that it is used to stabilize WM traces, but does not necessarily lead to stronger encoding into WM. Instead, the additional free-time is used to make a representation more robust to interference. This implies that in this mechanism, additional free-time will not necessarily lead to observable benefit to WM performance, unless the WM representation becomes degraded. It is important to note that the current study assumes stabilization and consolidation as two completely separate mechanisms. This implies that a weakly-encoded WM representation can nevertheless be stable and more resistant to interference. Conversely, an item which has been strongly encoded into WM can be less stable and less resistant to interference.

Free Time and Serial Position Curves

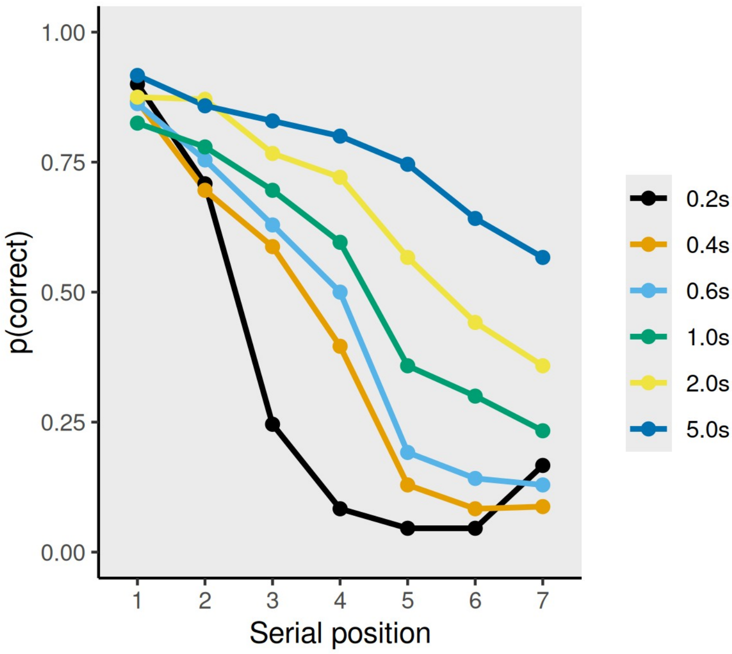
Recently, Oberauer (2022) parametrically manipulated the presentation rate of items for various materials (digits, letters, concrete words, abstract words…) and presentation modalities (visual, auditory), and he evaluated the impact of presentation rate on recall performance as a function of the serial position of the items. Results of this study showed that memory performance peaked for the first-encoded item and gradually declined across serial positions, a well-known phenomenon in immediate serial recall referred to as the *primacy effect*. Importantly, the free-time benefit was virtually absent for the very first item and gradually appeared and increased across serial positions, thus producing a "fanning-out" pattern. A portion of the Oberauer (2022) results' study is displayed in **Figure 1**[1]. Oberauer compared this pattern of results with predictions expected from different models, including a decay and refreshing/rehearsal model, a consolidation model, an encoding-resource model[2], and the temporal distinctiveness model (Brown et al., 2007). Results of these simulations showed that the encoding-resource best-aligned with the observed pattern of results for two main reasons. First, in this model, memory strength is maximal for the initial item and diminishes as memory resource is depleted, leading to a primacy effect. Second, slower presentation rates allow the resource to replenish more, thus leaving more resource available for the subsequent items, leading to stronger encoding strength as more and more items are encoded into WM comparatively to faster presentation rates. This property of the encoding-resource mechanism explains the fanning-out effect observed when manipulating presentation rate.

Figure 1

Experiment 1 from Oberauer (2022)

[1]We report results from the visual presentation using concrete words only, because this experiment best matches our own investigation and is likely to be the most appropriate to disentangle the mechanisms introduced above. Results from the other experiments closely match those illustrated in **Figure 1**. Readers interested by the results from the other experiments can read the original Oberauer (2022) study.

[2] Oberauer also considered a "ballistic consolidation" model. As Oberauer suggested that the ballistic consolidation and encoding-resource explanations could essentially represent two distinct interpretations of the same mechanism, and given their striking similarity in terms of predictions, we will consider only the encoding-resource mechanism.



Note. Experiments involved participants encoding and recalling (i.e., typed recall) lists of visually presented, concrete words. The free-time benefit was virtually absent over serial position 1, and gradually appeared over successive serial positions.

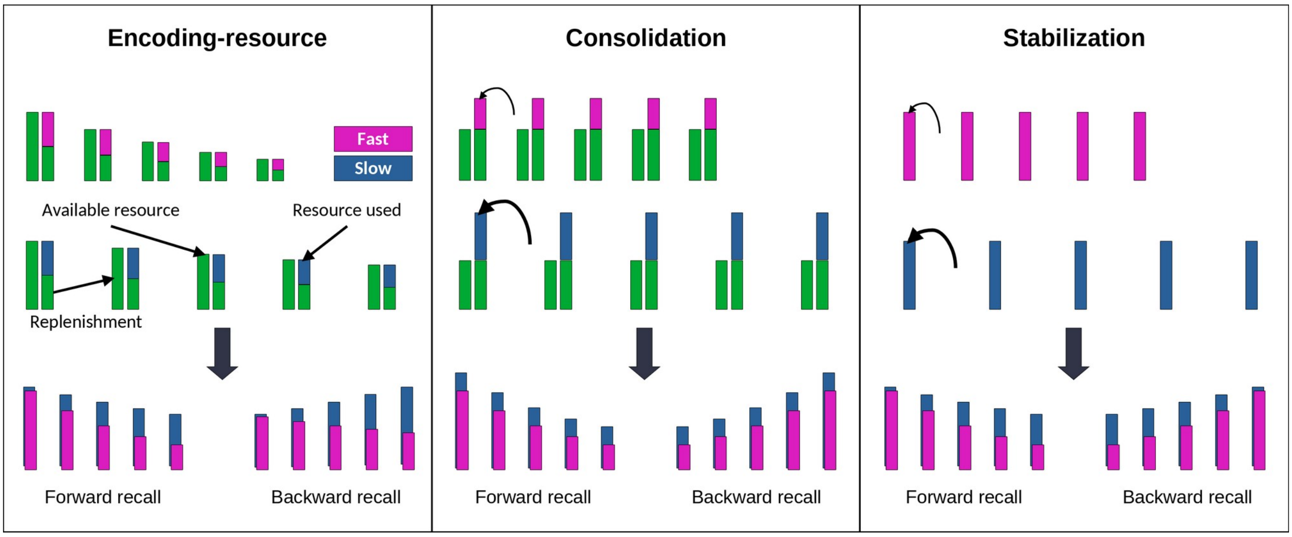
The above-mentioned study raises a number of questions. In comparing the models' predictions against empirical data, Oberauer assumed that the specific form of serial position recall curves only reflects encoding strength. However, it has been shown that the primacy effect in immediate serial recall stems from at least two phenomena. The first phenomenon can indeed be considered as a diminished encoding strength across serial positions, for example in the form of a primacy gradient of activation (Page & Norris, 1998) or through the encoding-resource mechanism as explained above. The second phenomenon is output interference, whereby recalling an item hinders subsequent items to be remembered (Cowan et al., 2002; Oberauer, 2003). For instance, when recalling “dog – desk – arm”, the mere fact of recalling “dog” degrades the representation of “desk” and “arm”. Therefore, items recalled later in the list are more likely to be poorly recalled by the mere fact that they suffer the most from output interference, and this phenomenon contributes to the primacy effect. In tasks requiring participants to recall items in their original presentation order, such as immediate serial recall, output position is fully confounded with input position, making it impossible to differentiate between effects occurring at encoding and those occurring at retrieval. One manifestation of output interference can be observed when participants recall items in backward (i.e., reverse) order. In this *backward recall* procedure, recall performance is best for the last encoded item, and progressively decreases until the first encoded item, for which recall performance is the worst (Dougherty et al., 2023; Guérard et al., 2012; Guérard & Saint-Aubin, 2012; Liu & Caplan, 2020). Similarly, when items are cued in a completely random order, the magnitude of the primacy effect diminishes and serial position curves are bow-shaped and symmetric (Kowialiewski, Krasnoff, et al., 2023; Oberauer, 2003). If items resist more strongly to output interference after being encoded for a longer period of time, for instance through the stabilization mechanism as described above, it is expected that output interference should cumulatively interfere less with the subsequent to-be-recalled items with additional free time, thus producing a less steep primacy effect with slower presentation rates. It is therefore possible that the fanning-out effect (i.e., serial position curves becoming increasingly less steep with increasing free time) observed by Oberauer (2022) stems from stronger resistance to output interference.

The Present Study

The current study adopts a comprehensive computational modeling and human data prediction strategy to evaluate the different mechanisms of the free-time benefit introduced above. We start by describing the WM mechanisms we use to predict the free-time benefit. We then compared the models' predictions against freshly acquired data and select the best models. To do this, we ran three experiments. These experiments involved participants recalling lists of concrete words in different recall orders. Recall order was post-cued in each experiment, thus ensuring that items were encoded in a similar way across the different recall conditions. In Experiment 1, we manipulated presentation rate in a forward-backward recall paradigm. Participants recalled items either in their original presentation order, or in reverse order. If the fanning-out effect observed by Oberauer (2022) is only caused by an encoding-resource mechanism, it should persist across input position when tested in a backward recall paradigm. In contrast, if the fanning-out effect is due to processes partially occurring during output, this effect should not be observed anymore as a function of input position but instead as a function of output position. Experiment 2 was a conceptual replication of Experiment 1, except that we introduced a distractor task between encoding and recall for reasons explained later. Finally, Experiment 3 used a cued recall paradigm in which items were tested randomly, which allowed us to further deconfound encoding- versus recall-based explanations of the free-time benefit. General predictions from the three proposed mechanisms are displayed in **Figure 2**.

Figure 2

Predictions from the three proposed mechanisms



Note. *Encoding resource*: Each trial starts with a fixed amount of resource, indicated by the green bars. Encoding is done by depleting a proportion of this resource, as indicated by the purple and blue bars for the fast and slow conditions, respectively. The amount of depleted resource determines encoding strength. During the free time available between two successive encoding operations, the resource replenishes. The slower the presentation rate, the more resource can replenish over time. This mechanism predicts a fanning-out effect increasing across input position, both in forward and backward recall. *Consolidation*: During the free time given between two successive encoding operations, the previously encoded item continues to be encoded. Additional free time increases encoding strength. This mechanism predicts that the free-time benefit should be uniformly observed across serial positions, both in forward and backward recall. *Stabilization*: After encoding, items are stabilized in such a way that they resist more strongly to output interference. Slower presentation rates result in stronger stabilization and therefore stronger resistance to output interference. This mechanism predicts a fanning-out effect following the direction in which items are recalled.

Computational Modeling

All the to-be-tested mechanisms presented here are integrated within a single architecture based on general principles that are commonly accepted in the WM literature. We first introduce these general principles, followed by the mechanisms responsible for modeling the free-time benefit.

Generic Mechanisms

Encoding. Encoding an item into WM is done by creating a new association between that item and a positional context (Farrell, 2012; Henson, 1998; Hitch et al., 1996; Oberauer et al., 2012; Oberauer & Lewandowsky, 2011). Positional contexts are similar to each other, and this similarity decreases exponentially with positional distance. This implies that items will be partially associated to other contexts than the one they were initially encountered (e.g., item 1 to position 2). The association strength $a_{i,j}$ between item i and position j follows this equation:

(1) $a_{ij} = \eta_i P^{|i-j|}$

Where P is a free parameter controlling the similarity between contexts. High values of this parameter mean high positional uncertainty. The η_i term is the encoding strength at input position i. This encoding strength depends on several mechanisms. One commonly held assumption is that encoding strength follows a primacy gradient of activation, whereby encoding strength decreases progressively for each newly encoded item (Page & Norris, 1998). We modeled this process as follow:

(2) $\eta_i = \alpha \gamma^{i-1}$

Where α is the peak activation of the primacy gradient, and γ controls the steepness of the primacy gradient. For simplicity, we assume that the first item is always maximally encoded, thus fixing the α parameter to 1.0. We varied the γ parameter freely.

Retrieval. When trying to retrieve an item, people need to use the cue which is currently available to them. In immediate serial recall, this cue is the position to which the item was associated to during encoding. For instance, when trying to retrieve the first item, people can use the cue “position 1”. However, due to the positional uncertainty P as described in Eq. 1, not only the target item will be re-activated following this cue, but also all other items associated to this positional cue, resulting in some uncertainty. Based on this cueing process, a pattern of activation is generated. Items are selected based on this pattern of activation, such that highly activated items have a higher probability to be selected than less activated items. Activation of list item i for output position k is:

(3) $A_{ik} = a_{ij} (1.0 - \rho)^{k-1}$

The ρ term is a free parameter which controls the strength of output interference. During the first retrieval attempt, $k-1=1-1=0$, which means that output interference doesn't have any effect yet, but grows from output position 2 on. At retrieval, the model has a certain probability to recall items that were not part of the list (i.e., extra-list intrusions) as well as omission errors. Theoretically, the production of extra-list intrusions is explained by assuming that non-list items have a certain degree of similarity with the retrieval candidates. This is implemented by giving them their own activation value: $A_{N+1} = \omega$. We fixed ω to 0.0, which when used in combination with the exponential version of Luce's choice rule as described below, gives extra-list intrusions a non-null probability to be selected. Omission errors are modeled using a threshold θ (free parameter). If an item's activation falls below this omission threshold, it is not recalled and an omission is produced. We implement this principle mathematically by giving the omission threshold an activation value and entering it into the competition: $A_{N+2} = \theta$.

After computing all items' activation given a recall cue, activations are converted into probabilities using the exponential version of Luce's choice rule (or softmax function):

$$(4)\ \ p_i = \frac{\exp\left(\frac{A_{ik}}{\sigma}\right)}{\sum_{j=1}^{N+2} \exp\left(\frac{A_{jk}}{\sigma}\right)}$$

In this equation, the σ parameter is the noise, and controls the steepness of the selection process. As this is an important parameter, it was estimated freely. High σ values means that all retrieval candidates become less distinctive to each other, thus increasing the probability that the selection process will select items randomly. When σ is low, activation values as computed in Eq. 3 are more deterministic of the selection process. This version of Luce's choice rule corresponds to a selection process in which items are selected based on their activation values, after adding normally distributed noise centered around 0.0, as typically done models of serial recall (Hurlstone & Hitch, 2015).

Modeling recall direction. One advantage of positional models is the fact that they provide sufficient flexibility for modelling any recall direction. All these models need for retrieving an item is a relevant cue. In the experiments reported in this study, the cue is the position of the item in the list. Therefore, when modeling a particular recall direction, we assume that participants use the cue which is currently available to them, and try to retrieve the item associated to it. In forward serial recall, the cues are the following: [1, 2, 3, 4, 5, 6]. In backward serial recall, the cues are: [6, 5, 4, 3, 2, 1]. Likewise, in cued recall, the cues can be any sequence of positional cues. Recall direction was therefore modeled by feeding the model with the same sequence of positional cues as the one given to our participants.

Response suppression. In many computational models of WM, items are discarded from the set of retrieval candidates after being recalled, a mechanism called response suppression (Lewandowsky, 1999). Response suppression is necessary in all models of WM requiring multiple recall attempts. Without it, WM models produce a rate of repetition errors which is unrealistic compared to what is observed in humans. However, recent studies have shown that people can recall multiple times the same item (Cowan & Hardman, 2021; Kowialiewski & Oberauer, 2024), an observation which contradicts a core prediction derived from response suppression. Despite solid doubts about the plausibility of response suppression as a fundamental mechanism of WM recall performance, there currently exists no other alternative to it. We therefore kept this response suppression mechanism, acknowledging that this choice should be regarded as a temporary solution to an yet unsolved problem. Response suppression is modeled by multiplying the recalled items' activation value by (1.0 - τ). The higher the value of τ, the stronger the suppression. We assume that response suppression is maximal (i.e., $\tau = 1.0$) to restrict the number of free parameters to a minimum.

After describing the general architecture used to model WM, we now describe in the next sections the mechanisms responsible for generating the free-time benefit.

Encoding Resource

The encoding resource mechanism follows the same principles as those reported in Mizrak and Oberauer (2021) as well as Oberauer (2022). At the beginning of each trial, the encoding resource R_1 is maximal (i.e., fixed to 1.0). During encoding, a constant proportion p_r of this resource is used as the encoding strength:

(5) $\eta_i = p_{r,i} R_i$

The amount of resource used for encoding is then depleted from the pool of available resource:

(6) $Rbis_i = \left(1.0 - p_{r,i}\right) R_i$

The p_r parameter was estimated freely. After encoding, the resource recovers with rate r_e:

(7) $R_{i+1} = Rbis_i + \left(1.0 - R_i\right)\left(1.0 - \exp\left(-r_e t_i\right)\right)$

Where t_i is the free-time available after the initial encoding. The r_e parameter was fixed to 0.44. This way, the encoding-resource and primacy gradient mechanisms can be compared based on an equivalent number of free parameters. Because the properties of the encoding-resource mechanism produce a primacy gradient of activation, we replaced Eq. 2 by Eq. 5 through 7 each time we considered this mechanism. The parameter value for r_e was chosen after performing a grid search over the encoding-resource mechanism's parameter space. Specifically, we orthogonally varied the p_r and r_e parameters and computed the mean free-time effect (expressed as the mean difference between two presentation rate conditions) this mechanism produces under each set of parameters.

With these computations, we found that fixing r_e to 0.44 while estimating p_r freely offers the most flexible range of possible outcomes (see **Appendix A** for details).

Consolidation

During consolidation, the just-encoded item continues to be re-encoded as a function of the total time t available to encode the item at a consolidation rate C, which was estimated freely. To compute the consolidation time of each item, we took back the equation reported by Oberauer (2022):

$$(7)\ \Delta\eta_i = \left(1.0 - \exp\left(-C t_i\right)\right)$$

The $\Delta\eta_i$ term means that consolidation adds an additional encoding strength on top of the existing one. For instance, if encoding strength for the first item is equal to 1.0, and the consolidation strength computed in Eq. 8 equals 0.2, the final encoding strength is equal to 1.2.

Stabilization

During stabilization, items are consolidated in such a way that they resist more strongly to interference. The role of this consolidation mechanism is not to increase items' encoding strength, but rather to stabilize the existing WM representation. Due to this stabilization process, items suffer less from output interference. Adapting Eq. 3, it gives:

$$(8)\ A_{i,k} = a_{i,j}\left(1.0 - \rho\right)^{(k-1)\lambda/t_i}$$

In this equation, λ is a free parameter which controls the overall reduction of output interference with free time. The t_i term is the time spent consolidating item i.

Model fitting

The data collected in Experiments 1 through 3 were fitted to the above-mentioned mechanisms. To do this, we performed a quantitative fit of the different models using maximum likelihood estimators. This way, we can perform model comparison and select models based on their quantitative fit to the data. Fixed and free parameters are reported in **Table 1**.

Table 1. List of fixed and free parameters of the model.

Symbol	Role	Value
	Core WM architecture	
P	Positional overlap	[0.1 – 0.8]
α	Peak of the primacy gradient	1.0
γ	Steepness of the primacy gradient	[0.0 – 1.0]
ρ	Output interference	[0.0 – 1.0]
ω	Activation of the non-target items	0.0
θ	Value of the omission threshold	[0.0 – 10.0]
σ	Noise parameter used during the selection rule	[0.0 – 1.0]
τ	Response suppression	1.0
	Free-time mechanisms	
	Encoding resource mechanism	
R_1	Initial resource	1.0
p_r	Proportion of resource used	[0.1 – 0.9]
r_e	Rate of resource replenishment	0.44
	Consolidation mechanism	

C	Consolidation rate	[0.0 – 10.0]
	Stabilization mechanism	
λ	Reduction of output interference	[0.0 – 5.0]

Note. Fixed parameters are indicated by a single value. Free parameters are indicated by a range.

Fitting procedure. Model fitting was done at the subject level using the raw data (i.e., non-aggregated individual trials). For each recall attempt, we computed the probability to recall each retrieval candidate using Eq. 4. Based on this recall probability, we computed the log-likelihood for the observed response o:

(10) $logL = \log(p_o)$

Note that in this fitting procedure, when applying response suppression, we used the *observed response* produced by the participant. We used the deviance as a loss function:

(11) $D = -2.0 \sum logL$

In Eq 11., the sum operator runs over all trials and retrieval attempt.

Parameter estimation was done using the Nelder-Mead algorithm implemented in the Optim package (https://julianlsolvers.github.io/Optim.jl/stable/) of the Julia programming language (https://julialang.org/benchmarks/). To avoid that the algorithm would fall into local minima, each fitting attempt was repeated using 15 different starting points in the multi-dimensional parameter space. These starting points were randomly selected by sampling values from a uniform distribution. We kept only the set of parameters minimizing the deviance.

Model comparison. To compare models with each other, we first computed a Bayesian Information Criterion (BIC) for each model:

(12) $BIC = \sum_{i=1}^{N} K \log(n_o) + D_i$

In Eq. 12, the sum runs over all participants. We therefore computed one BIC for each participant, and summed all BICs to get an overall assessment for a particular model. The K term is the number of free parameters, and n_o is the number of observations per participant. In Experiment 1, this number is equal to 6 items, times 64 trials. We chose the BIC because it penalizes more strongly models with a larger number of free parameters. After computing one BIC for each model, models can be compared by subtracting their BIC. This difference represents the likelihood of the data under a certain model relatively to another model.

Experiment 1

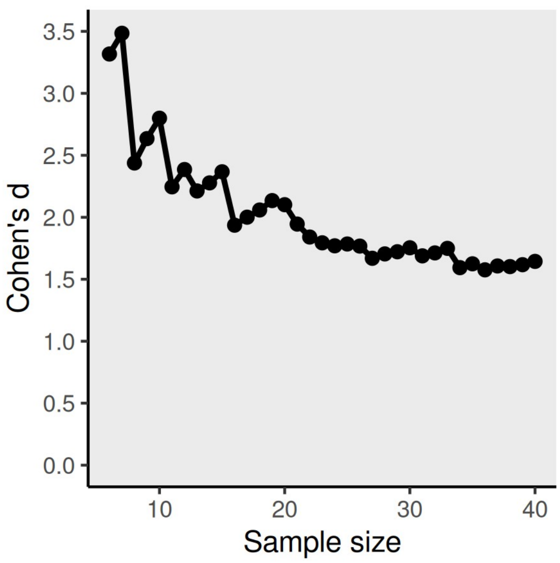
This experiment manipulates presentation rate across two different recall conditions: a forward recall condition in which participants recall items in their original presentation order, and a backward recall condition in which participants recall items in *reverse order*. The stimuli involved lists of semantically dissimilar, concrete words. We chose concrete words because they are more likely to be consolidated than any other class of stimuli, which maximizes the chances to observe an influence of consolidation, if any. In addition, results from Oberauer (2022) showed that lists of concrete and abstract words were differently affected by the presentation rate manipulation, suggesting a possible consolidation process for concrete words.

Methods

Optional stopping. Sample size was determined using an optional stopping rule based on effect size stabilization (Anderson et al., 2022). We chose this way of defining our sample size because simulations have shown that it does not inflate effect sizes nor increase the rate of false positives, contrary to other stopping rules. We started with a base sample size of 15 participants. We then iteratively added one more participant to the sample and computed the effect size (Cohen's d) each time a new participant entered the sample. If the difference between effect sizes did not exceed 0.05 over five successive iterations, we stopped the sampling process, which determined our final sample. For instance, if effect sizes reached values of 0.8, 0.83, 0.82, 0.78 and 0.81 after 32, 33, 34, 35 and 36 participants, the sampling process stopped at a sample size of 36 participants. Using this procedure, we reached a sample size of 40 participants. The effect size was computed based on the difference between the fastest and the slowest presentation rate conditions (see procedure below). We report in **Figure 3** a graph illustrating the evolution of effect sizes over the sampling process. As can be seen, the effect size gravitated at d ~ 1.5.

Figure 3

Effect size stabilization – Experiment 1



Note. Cohen's ds were computed using the difference between the fastest (0.5 seconds / item) and slowest (4.0 seconds / item) presentation rate conditions.

Participants. Forty young adults aged between 18 and 35 years were recruited on the online platform Prolific (https://prolific.co/). All participants were English native speakers, reported no history of neurological disorder or learning difficulty, and gave their written informed consent before starting the experiment. The experiment had been approved by the ethics committee of the Faculty of Psychology at the University of Liège, project #2021-024.

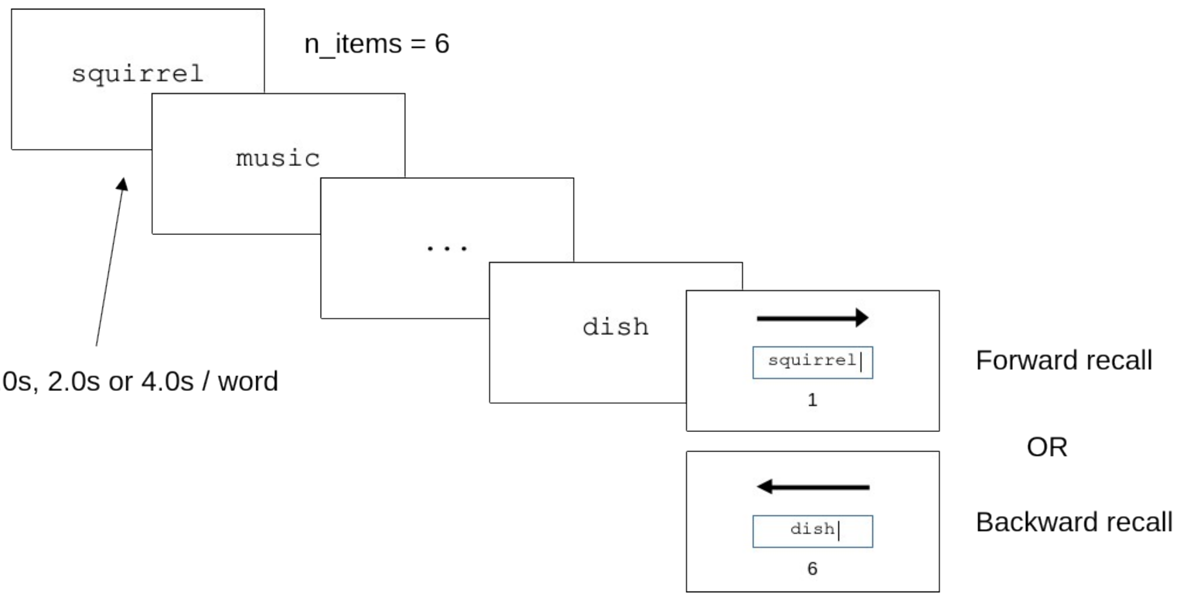
Design. Participants were tested across eight experimental conditions in a fully within-subject design: two recall conditions (forward, backward) times four presentation rate conditions (0.5, 1.0, 2.0 and 4.0 seconds per item). Each experimental condition involved 8 trials, for a total of 64 experimental trials. We chose this number of trials to keep the experiment at a reasonable length.

Material. The list of stimuli was constructed from a pool of 312 concrete words. The initial pool consisted in 39 categories of 8 words. To construct the lists, we randomly chose six words among different categories, with the further constraint that two words could not be included in the same list if they were drawn from two related or similar categories. For instance, words "whiskey" and "glass" could never appear together in the same list, because "alcohol" and "containers" are two strongly related categories. We constructed the lists this way to avoid any obvious semantic relationship or similarity between items, for which the interaction with free time is still unknown. This is a classical way to construct dissimilar lists in studies manipulating semantic similarity (Kowialiewski, Krasnoff, et al., 2023; Kowialiewski, Majerus, et al., 2023; Neath et al., 2022; Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999). We constructed 32 lists of 6 items twice: once for the forward recall condition, and once for the backward recall condition. The lists in the forward and backward conditions were constructed by sampling from the same pool of 312 concrete words. This means that some words may have been presented twice across the whole experiment: once in the forward recall condition, and once in the backward recall condition. Within each recall condition, there were 8 trials per presentation rate condition. Each list was assigned randomly to a presentation rate condition. The order of the presentation rate and recall conditions was random. There were 4 training trials, for which the lists of stimuli were always identical across participants and generated in advance. The training and main phases of the experiment always involved different stimuli. Overall, there was a total of 68 trials throughout the experiment: 4 training trials, and 64 experimental trials. Using these aforementioned constraints, we generated 120 different versions of the experiment. Each participant was assigned randomly to one of these versions.

Procedure. The task is illustrated in **Figure 4**. Words were sequentially presented in the center of the screen in Courier font, for a duration defined by the presentation rate condition (i.e., 0.5, 1.0, 2.0 or 4.0 seconds). Each word remained on screen until the presentation of the next word. After the presentation of the last word, an arrow indicated the direction of retrieval. If the arrow pointed to the right, participants were instructed to recall items in forward order. If the arrow pointed to the left, participants were instructed to recall items in backward order. Participants entered their responses in a prompt box using the keyboard of their computer and validated each response using the “return” key. If participants did not know the answer for a given position, they were instructed to leave the prompt box empty. As participants recalled the items, a number below the prompt box was displayed to indicate the position of the current to-be-recalled word. After recalling the last item, participants clicked on a button labeled “Next trial” to move on to the next experimental trial. Halfway through the experiment, participants could take a short break if they needed to. Throughout the whole experiment, recall and presentation rate conditions were post-cued. This means that participants never knew in advance the experimental condition they had to perform on a particular trial. During the training phase, participants performed two forward recall conditions and two backward recall conditions, in this order. During training, words were presented at a pace of 1 word every second. After the training phase, participants were warned that the presentation rate of words would vary randomly from one trial to another.

Figure 4

Setup used in Experiment 1



Note. Participants were visually presented with six items to be remembered. Presentation rate varied depending on the condition (i.e., 0.5, 1.0, 2.0, or 4.0 seconds per item). After the presentation of the last word to be remembered, participants were presented with an arrow, along with a prompt box. Participants had to recall the words in forward (i.e., arrow pointing to the right) or backward (arrow pointing to the left) serial order.

Scoring procedure. Before scoring responses as correct or incorrect, we preprocessed them by removing blank spaces and transforming uppercase letters to lowercase. We used a strict serial recall criterion in which an item was scored as correct if it was recalled at the correct serial position.

Statistical analyses. As the focus of the manuscript was to compare the models' outcome against the empirical data, we kept the statistical analyses to a bare minimum. We assessed the effect of presentation rate on recall performance with a Bayesian logistic regression model using the brms package (Bürkner, 2017), assuming default priors. Each model was run using 4 chains of 10,000 iterations, including 5,000 warm-up iterations. To get the strength of evidence for a particular effect, we performed Bayesian model comparison using a top-down testing procedure. We assessed each effect of interest by comparing the full model to the same model without the effect in question using the bayes_factor() function provided in the brms package.

Results

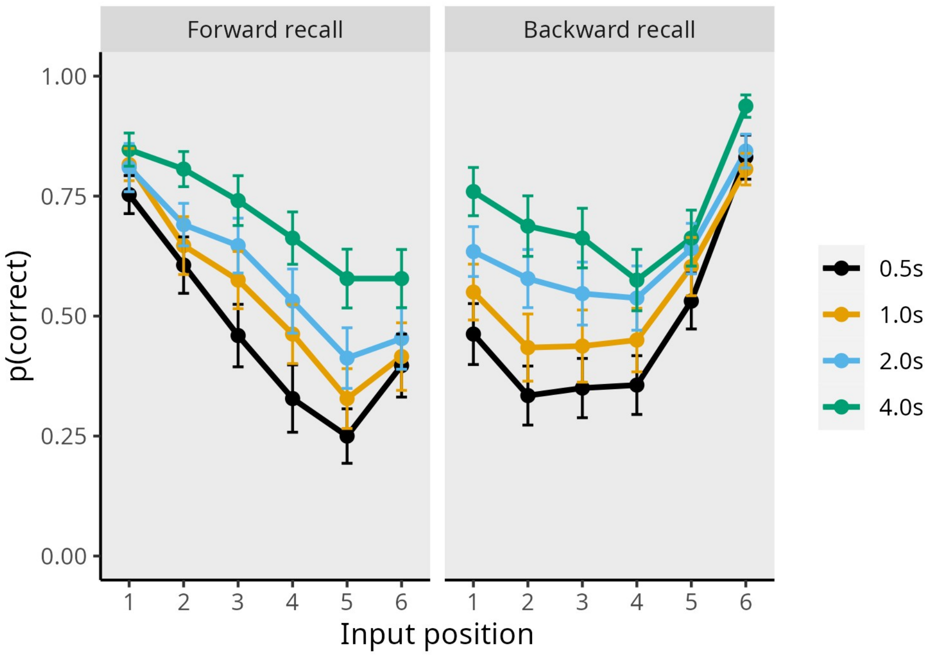
Figure 5 shows recall performance as a function of presentation rate, input position and recall direction. Memory performance increased with presentation rate, and this was observed across both recall directions. The Bayesian logistic regression model showed decisive evidence supporting an effect of presentation rate, both in the forward (BF_{10} = 3.19e+8) and backward (BF_{10} = 1.125e+10) recall directions.

As can be seen in **Figure 5**, the free-time benefit increased across input position in the forward recall direction. In contrast, the free-time benefit decreased across input position in the backward recall direction[3]. A visual inspection of the forward recall direction suggests that the free-time benefit was already apparent from input position 1, which was confirmed by a logistic regression model (BF_{10} = 19).

Figure 5

Empirical results – Experiment 1

[3]A reviewer wondered if a "peel-off" strategy might have been used by participants in the backward recall condition, which should be characterized by longer response times across output position. We did not observe such a pattern of results. This analysis has been made available on the OSF repository associated with this study, for Experiments 1 and 2.



Note. Error bars indicate 95% confidence intervals, corrected for within-subject variability.

Modeling section

We fitted the generic model using the three mechanisms we identified: consolidation, stabilization, and encoding-resource. When considering the encoding-resource mechanism, we omitted the primacy gradient of activation to avoid redundancy, as the encoding-resource mechanism produces a primacy gradient of activation. The best model was the model including the stabilization mechanism. Subtracting BIC values between models showed that the data were 163 times more likely under the stabilization model than under the encoding-resource model, and 94 times more likely under the stabilization than the consolidation model. A summary of this model comparison approach is reported in **Table 2**.

Table 2. Model comparison from Experiment 1	
Models	*BIC*
[primacy gradient] + [stabilization]	34,817

[primacy gradient] + [consolidation]	34,911
[encoding resource]	34,981
Note. Low BIC values indicate better fit to the data.	

Predictions from the three models, along with the empirical data, are displayed in **Figure 6**. All models predict a free-time benefit on recall performance. In the following paragraphs, we briefly describe the behavior of each model, and explain why they account more or less well for the empirical data.

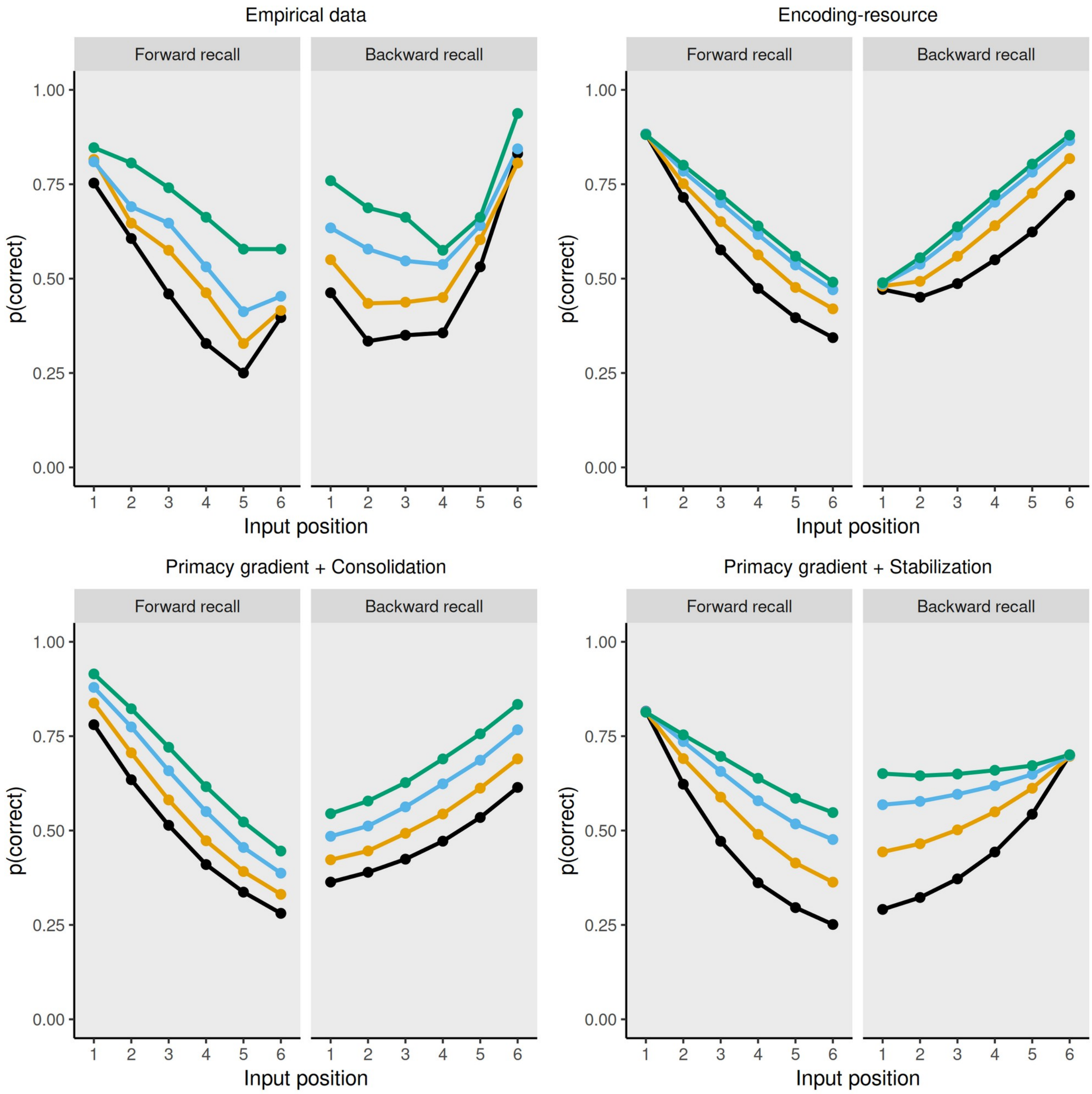
Encoding resource. **Figure 6** shows that the encoding-resource mechanism predicts better recall performance with increasing free time, as already shown in previous works (Mizrak & Oberauer, 2021; Oberauer, 2022). Because the free-time benefit generated by this mechanism builds-up progressively at encoding, this benefit gets stronger across input position, regardless of the recall condition considered (i.e., forward recall, backward recall). This is also the reason why this mechanism gives the worst goodness of fit: In the empirical data, the free-time benefit increased across *output position*.

Consolidation. The consolidation mechanism also produces a free-time benefit, as originally shown by Oberauer (2022). Consolidation produces a constant free-time benefit across input position, because the consolidation mechanism considered here increases encoding strength to a similar extent across all items. Contrary to the encoding-resource mechanism, consolidation does not underpredict the data, as a constant free-time benefit across serial positions can still capture part of the variance observed in humans, including a free-time benefit for items encoded first (see Results section).

Stabilization. As expected, the stabilization mechanism produces nearly symmetrical serial position curves across recall directions. While the free-time benefit increases across input position in the forward recall condition, it decreased in the backward recall condition. In other words, this mechanism produces stronger free-time benefits across *output position*, in line with the data. This occurs because for each item recalled, the subsequent to-be-recalled items suffer less from output interference with slower presentation time, and this phenomenon builds up cumulatively each time the model recalls an item. Note that this mechanism never produces a free-time benefit from output position 1, as no output interference occurs for items recalled first.

Figure 6

Simulation results of Experiment 1 using individual free-time mechanisms



Note. The upper left panel displays the empirical data. All other panels display model predictions.

Additional Simulations

Results from the previous section support a stabilization mechanism whereby items become more resistant to output interference with slower presentation rate. It is however possible that the free-time benefit emerges from more than one mechanism. This is suggested by the behavioral results showing a free-time benefit for the very first encoded item in forward recall, a result which can be accounted only by the consolidation mechanism. In this section, we test this possibility. We used the generic architecture as presented above. Rather than considering models with only one free-time mechanism, we considered multiple mechanisms operating in concert. For instance, we considered a model including the consolidation, stabilization, and encoding-resource mechanisms, or a model including the consolidation and stabilization mechanisms only. We adopted a top-down approach as used in the statistical procedure described above. We started by fitting the most complex model or models. We then fitted another version of this model, for which a mechanism was removed or replaced. After selecting the model with the lowest BIC, we repeated the operation until reaching the best fitting model among all alternatives.

A summary of model comparison results is provided in **Table 3**. The first model we considered is a model including all three mechanisms responsible to produce the free-time benefit: consolidation, stabilization, and encoding-resource. This model was first compared to a model including consolidation, stabilization, and the primacy gradient of activation. These two models have an equivalent number of free parameters. This comparison led to a BIC difference of 138 in favor of the model including the primacy gradient of activation. Predictions from these two models are provided in **Figure 7**. As can be seen, both models make very similar qualitative predictions. This suggests that the encoding-resource mechanism did not substantially contribute to the free-time benefit in our dataset, beyond a primacy gradient of activation.

Next, we compared the model including consolidation, stabilization, and the primacy gradient of activation mechanisms, against the same model without the consolidation mechanism. The BIC difference shows that the data were 145 times more likely under the model not including the consolidation mechanism than under a model including it. **Figure 8** displays simulations results from the best model with or without the main mechanisms of interest. As can be seen, the model without the consolidation mechanism (i.e., Primacy gradient + Stabilization, upper right panel) does not predict the free-time benefit observed from input position 1. At the same time, the model with consolidation (bottom left panel) was most likely not favoured because it adds an extra free-parameter, which the BIC penalizes more[4]. Overall, the consdolidation mechanism is not necessary to capture the overall pattern of results. With this model being identified, we are back to the conclusions we reached in the previous section (i.e., "Modeling section", see also **Table 2**).

We performed a last model comparison involving the model including the stabilization and primacy gradient of activation mechanisms, against a model not including the primacy gradient. The BIC difference between these two models showed that the data were 717 times more likely under the model including the primacy gradient than under the model not including it. The empirical data in **Figure 8** shows that in the backward recall condition, items encoded first are better recalled than items recalled last in the forward recall condition. Without the primacy gradient of activation, this pattern wouldn't emerge and serial position curves in the forward and backward recall conditions would be fully symmetrical. An illustration of what would happen in this scenario is illustrated in **Figure 8**, bottom right panel ("Without primacy" model). Therefore, the best-fitting model is the model including the stabilization and primacy gradient of activation mechanisms. Predictions of this best-fitting model are displayed in **Figure 8**, upper right panel.

Table 3. Model comparison from Experiment 1

	Model 1	Model 2	$BIC_{m1} - BIC_{m2}$
Step 1	[consolidation] + [stabilization] + [encoding-resource]	[consolidation] + [stabilization] + [primacy gradient]	-138
Step 2	[consolidation] + [stabilization] +	[primacy gradient] +	145

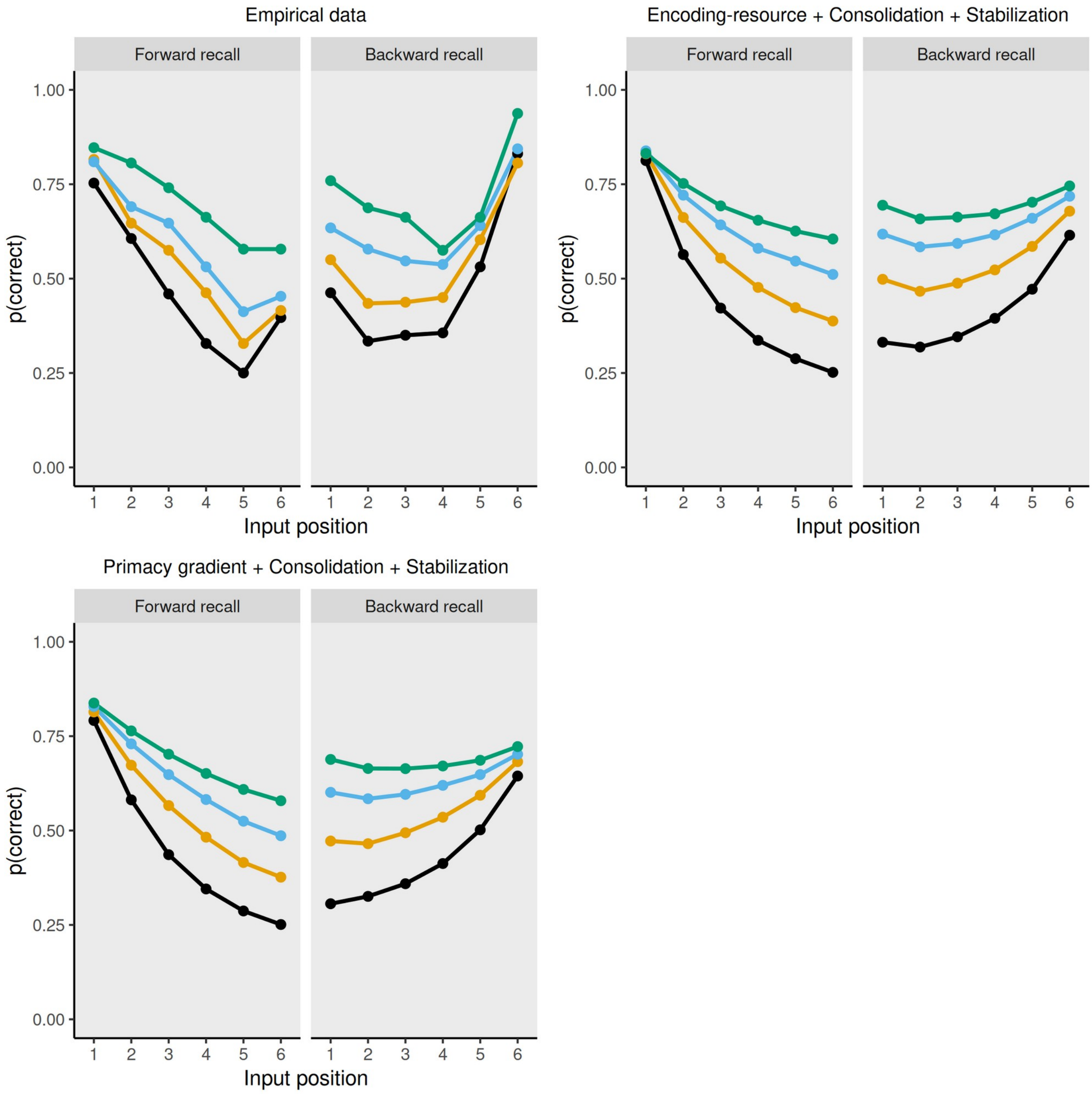
[4]Confirming this, using Akaike's Information Criterion (AIC) which penalizes less for model complexity favored the model including the consolidation mechanism.

	[primacy gradient]	[stabilization]	
Step 3	[primacy gradient] + [stabilization]	[stabilization]	-716

Note. Model comparison was performed using a top-down approach, by considering first the most complex models justified by our experimental design and theories. Negative values indicate better fit for m1 as compared to m2.

Figure 7

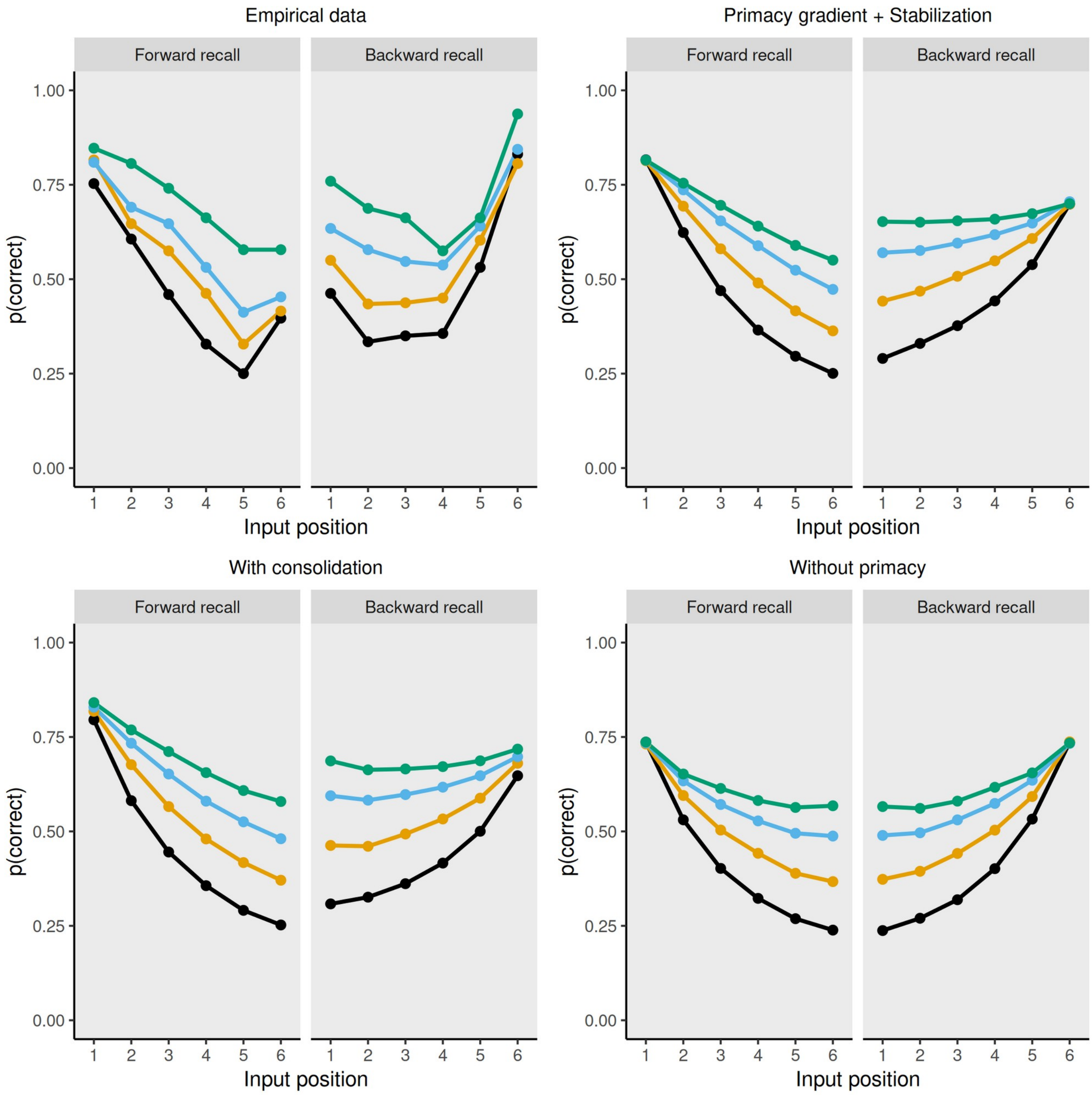
Model predictions from Experiment 1



Note. Upper left panel: empirical data. Upper right panel: model including a consolidation, stabilization and encoding-source mechanisms. Bottom left panel: model including a consolidation, stabilization and primacy gradient mechanisms. As can be seen, both models make very similar predictions.

Figure 8

540 *Simulation results from Experiment 1 – Best model with or without the main mechanisms*



541 *Note*. Upper left panel: Empirical data. Upper right panel: Best model. Bottom left panel: Best
542 model with the consolidation mechanism. Bottom right panel: Best model without the primacy
543 gradient mechanism.

544

Discussion

Results of Experiment 1 replicate the beneficial effect of presentation rate on memory performance. The novel aspect of this experiment was to show that the free-time benefit increased across output position. These results are well-accounted by a stabilization mechanism whereby extra free-time leads to more resistance to output interference. We also observed a credible effect of free time on the very first item in the forward recall condition, a result only predicted by a consolidation mechanism.

Additional simulations confirmed these preliminary observations to some extent. Model comparison clearly favored the stabilization mechanism, which explains the ‘fanning-out’ effect observed both in forward and backward recall. However, the consolidation mechanism was not favored, which contrasts with the observation of a free-time benefit for the very first item in the forward recall condition which none of the other mechanisms can explain. Note that a primacy gradient of activation was favored compared to an encoding-resource mechanism. The encoding-resource mechanism did not substantially change the overall qualitative pattern of results compared to a primacy gradient mechanism. Model fit was worse when not including the primacy gradient of activation, which corroborates with previous modeling works (e.g., Hurlstone & Hitch, 2015).

The fanning-out effect observed in backward recall might have been exaggerated by the fact that in backward recall, the last encoded item is also recalled first. This represents a potential limitation, because it is likely that the last item was still strongly represented in WM at the time of recall, for instance by virtue of being in the focus of attention (Cowan, 1999; Oberauer, 2002), which is not implemented in our simulations. If this item is in the focus of attention, it should be very strongly represented in WM and might therefore not benefit from slower presentation rates. It is thus possible that the fanning out effect in our empirical results is not only to be explained by direction of recall, but also by the specific status of the last encoded item. A closer examination of

Figure 5 indicates that this is likely to be the case, as recall performance for the last encoded item in the backward recall condition was up to ~80% in the fastest condition, and dropped to nearly 50% for the penultimate item. To appropriately estimate the presence of a potential free-time mechanism, it is therefore important to discard an influence from the last-encoded item. Experiment 2 addresses this directly.

Experiment 2

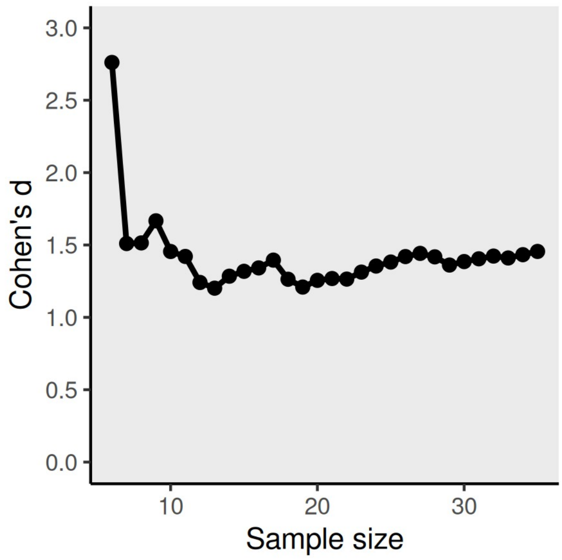
Experiment 2 replicates Experiment 1, except that we introduced a distractor task between encoding and recall. If the strong fanning-out effect observed across *output position* (rather than input position) observed in Experiment 1 is partially caused by the last item being in the focus of attention, the distractor task should prevent that item from being too strongly represented at the time of retrieval. This should in turn mitigate the strong fanning-out effect across output position observed in Experiment 1. If the strong fanning-out effect across output position is not due to the last item being in the focus of attention, we should replicate Experiment 1, except that recall performance should be comparatively lower due to the distractor task.

Methods

Optional stopping rule. Experiment 2 used the same optional stopping rule based on effect size stabilization as in Experiment 1. We reached a final sample size of 35 participants. The evolution of the effect size across sample size is displayed in **Figure 9**. Again, the effect size gravitated at $d \sim 1.5$.

Figure 9

Effect size stabilization – Experiment 2



Note. Cohen's ds were computed using the difference between the fastest (0.5 seconds / item) and slowest (4.0 seconds / item) time conditions.

Participants. Thirty-five young adults aged between 18 and 35 years were recruited on the online platform Prolific (https://prolific.co/). All participants were English native speakers, reported no history of neurological disorder or learning difficulty, and gave their written informed consent before starting the experiment. The experiment had been approved by the ethics committee of the Faculty of Psychology at the University of Liège, project #2021-024.

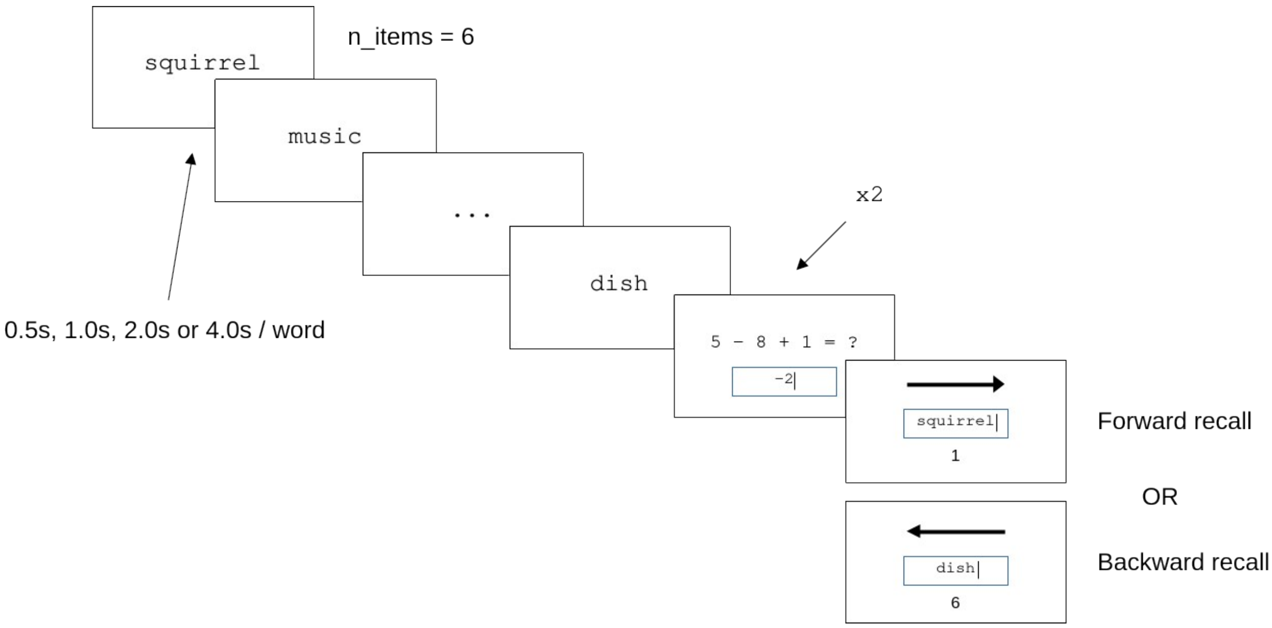
Design. The design of the experiment remained the same as Experiment 1, including the number of trials per participant.

Material. We used the same material as described in Experiment 1.

Procedure. The procedure was identical to Experiment 1, with one exception. After the presentation of the last item to be remembered, participants performed two rounds of mathematical equations. The equations involved three digits between 0 and 9, selected at random. We then randomly selected two mathematical operators, involving addition or subtraction. The mathematical operators were intersected between the digits and participants had to to solve the resulting equation by entering their answer in a prompt box. For instance, given the digits 5, 8 and 1, participants could be presented with the equation “5 – 8 + 1”, to which they had to answer “-2”. After resolving a first equation, participants were directly presented with a second equation. Providing a response to this second equation led to the presentation of an arrow pointing to the right of left, probing participants to recall the words in forward or backward order, as in Experiment 1. The experimental procedure is illustrated in **Figure 10**.

Figure 10

Setup used in Experiment 2



Note. Participants were visually presented with six items to be remembered. Presentation rate varied depending on the time condition (i.e., 0.5, 1.0, 2.0, or 4.0 seconds per item). After the presentation of the last word to be remembered, participants had to complete two mathematical operations in a row. Completing the two mathematical operations led to the presentation of an arrow, along with a prompt box. Participants had to recall the words in forward (i.e., arrow pointing to the right) or backward (arrow pointing to the left) serial order.

Scoring procedure. Participants' responses were scored as in Experiment 1.

Statistical analyses. Statistical analyses were conducted using the same strategy as in Experiment 1.

Results

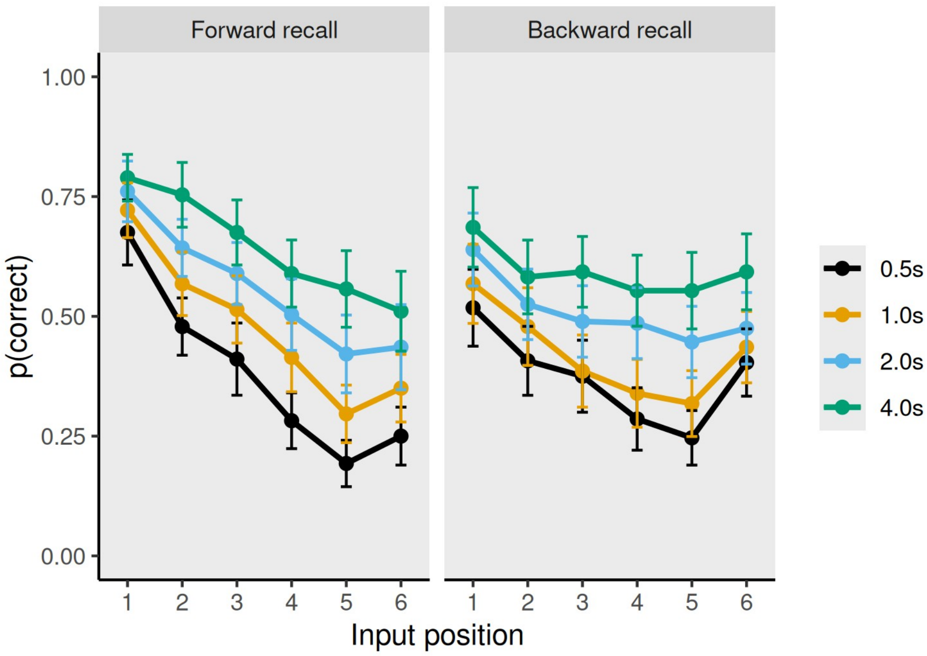
Recall performance as a function of presentation rate, input position and recall direction is displayed in **Figure 11**. A Bayesian logistic regression model indicates that memory performance increased with slower presentation time, and this was supported by decisive evidence both in the forward (BF_{10} = 4.802e+5) and backward (BF_{10} = 4,695) recall directions.

We furthermore explored to what extent a free-time benefit for the very first item in the forward recall condition could be observed, a result only predicted by the consolidation mechanism. A Bayesian logistic regression model with presentation rate as independent variable showed decisive evidence in favor of an effect of presentation rate (BF_{10} = 122).

We next explored whether a free-time benefit could be observed for the last item in the backward recall condition. A free-time benefit for this position would mean that the data cannot be exclusively explained by the stabilization mechanism, and therefore the direction of retrieval. A Bayesian logistic regression model showed strong evidence supporting an effect of presentation rate (BF_{10} = 42).

Figure 11

Empirical results – Experiment 2



Note. Error bars indicate 95% confidence intervals, corrected for within-subject variability.

Model fitting

Simulations in Experiment 2 followed exactly the same modelling procedures and steps as in Experiment 1, via fitting of parameter values. We did not explicitly model the arithmetic interfering task, because we considered that direct modeling of this task would add unnecessary complexity to the simulations. Instead, the impact of the interfering task was modelled via a change in fitted parameter values. This modeling section directly considered multiple free-time mechanisms operating altogether. We fitted the different mechanisms to the empirical data, and selected the best model based on BIC difference using a top-down approach. A summary of this model comparison analysis is provided in **Table 4**.

We started by considering a model including the consolidation mechanism, the stabilization mechanism, and an encoding-resource mechanism. We compared this model to the same model but by replacing the encoding-resource mechanism by the primacy gradient mechanism. As in Experiment 1, the data were 598 times more likely under the model including the primacy gradient of activation than the model including the encoding-resource mechanism. The reason why the encoding-resource mechanism did not provide a good fit of the data is clear when examining **Figure 12**. First, the encoding-resource mechanism produces an exaggerated fanning-out effect in the forward recall condition compared to what is observed in human subjects. Second, it also produces symmetrical serial position curves in the backward recall condition, and this is especially true when considering the slowest time condition. In contrast to this, serial position curves in the backward recall direction display a strong primacy effect, a pattern of results which is accounted only by the model including the primacy gradient mechanism. It is also important to note that the fanning-out effect is not as strong for the last-encoded item, which is also the reason why the encoding-resource model was not favored, as the free-time benefit necessarily increases for items encoded last. The stabilization mechanism does produce this reduced benefit for the last-encoded item, which is likely why it best-fitted the data.

Next, we compared the model including the consolidation, stabilization and primacy gradient mechanisms against the same model but without the primacy gradient. BIC comparison between these two models indicates that the data were 820 times more likely under the model including the primacy gradient mechanism than under the model without. The primacy gradient was favored over a model not including it, for the same reasons as those observed in Experiment 1: Without it, the model produces perfectly symmetrical serial position curves, which is not observed in the empirical data. Model predictions without this mechanism are illustrated in **Figure 13**, under the "Without primacy" panel.

Using the same approach, we compared the model including the consolidation, stabilization, and primacy gradient mechanisms to the same model without the stabilization mechanism. We found that the data were 296 times more likely under the model including the stabilization mechanism than the model not including it. Without this mechanism, the model produces an equivalent free-time benefit across serial positions, which is clearly in contradiction with the empirical data. Model predictions without the stabilization mechanism are reported in **Figure 13**, “Without stabilization” panel.

Finally, we compared the model including the consolidation, stabilization and primacy gradient mechanisms against the same model without the consolidation mechanism. Results show that the data were 153 times more likely under the model with the consolidation mechanism than the model without. The model without the consolidation mechanism predicted no free-time benefit from output position 1 (see **Figure 13**, under the “Without consolidation” panel), a result also in contradiction with the empirical data. Therefore, the best-fitting model was the model including the consolidation mechanism, the stabilization mechanism, and the primacy gradient mechanism. **Figure 12**, bottom panel, displays simulation results from this model. Goodness-of-fit for the best-fitting model is well-represented by looking at **Figure 13**, where we removed each mechanism from it. Without this specific combination of mechanisms, none of the models considered can match closely the pattern of empirical data.

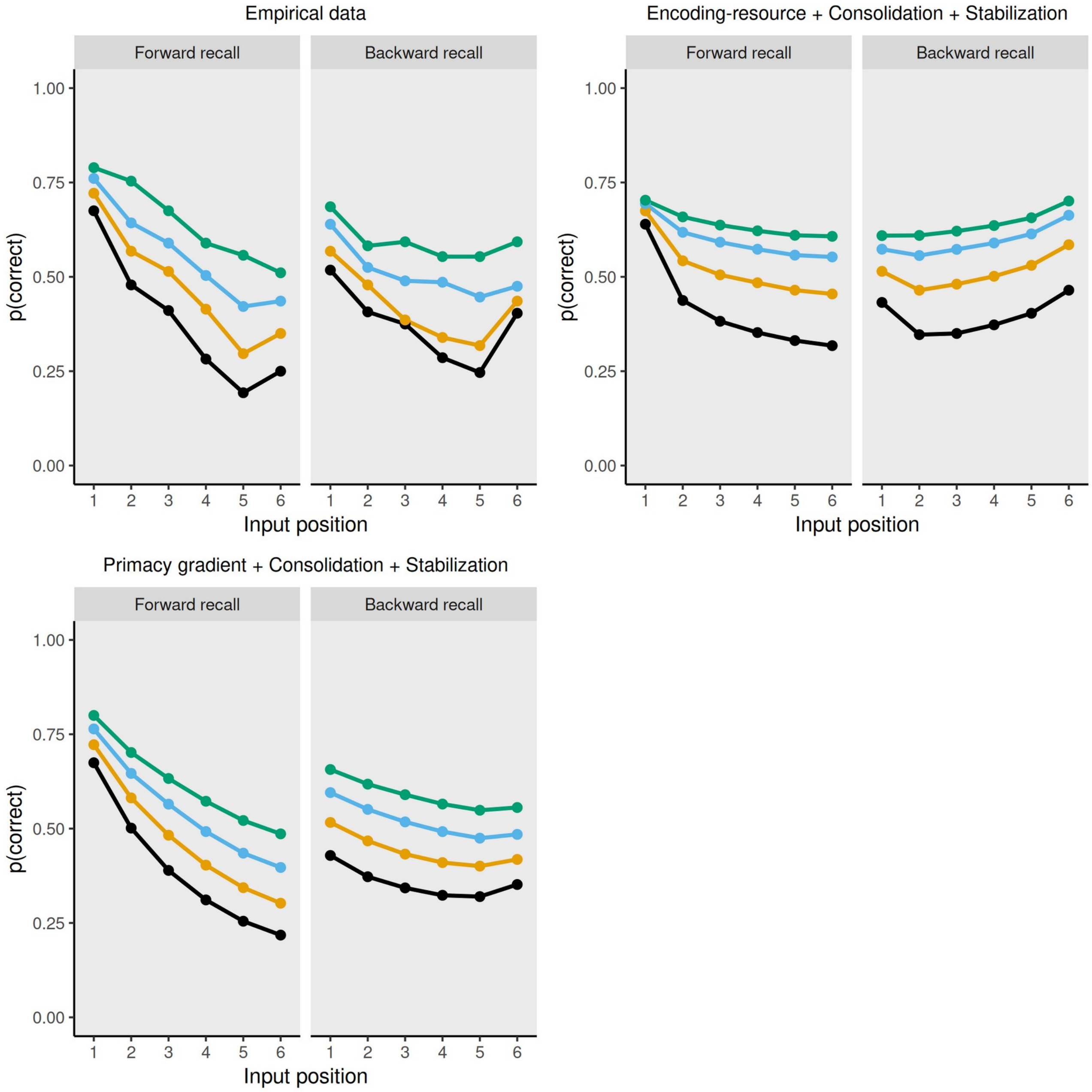
Table 4. Model comparison from Experiment 2

	Model 1	Model 2	$BIC_{m1} - BIC_{m2}$
Step 1	[consolidation] + [stabilization] + [encoding-resource]	[consolidation] + [stabilization] + [primacy gradient]	598
Step 2	[consolidation] + [stabilization] + [primacy gradient]	[consolidation] + [stabilization]	-820
Step 3	[consolidation] + [stabilization] + [primacy gradient]	[consolidation] + [primacy gradient]	-296
Step 4	[consolidation] + [stabilization] + [primacy gradient]	[stabilization] + [primacy gradient]	-153

Note. Model comparison was performed using a top-down approach, by considering first the most complex models justified by our experimental design and theories. Negative values indicate better fit for m1 as compared to m2.

Figure 12

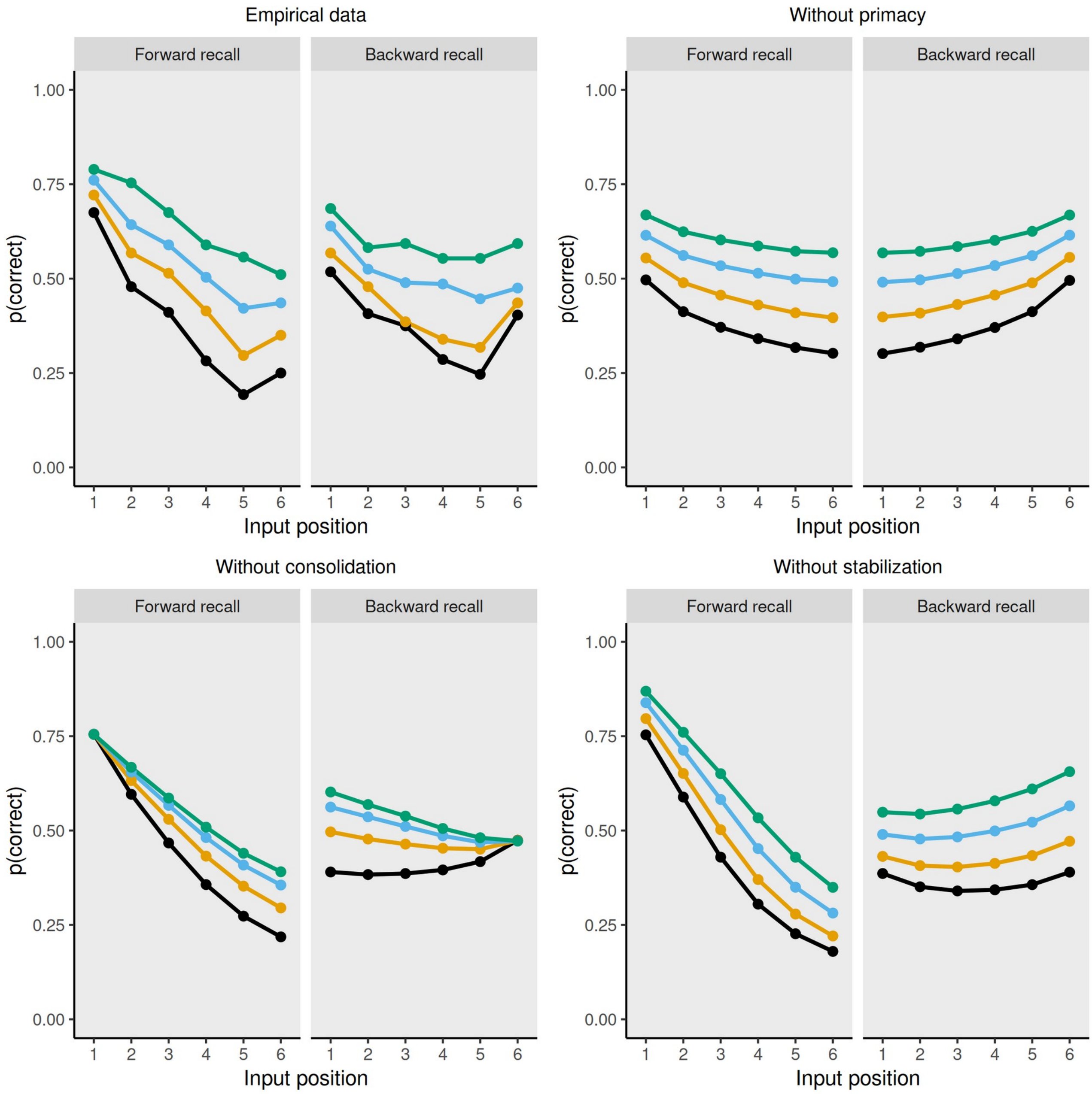
Model predictions from Experiment 2



Note. Top left panel: Empirical data. Top right panel: model including a consolidation, stabilization and encoding-resource mechanisms. Bottom left panel: model including a consolidation, stabilization and primacy gradient mechanisms. The winning model across the whole model comparison analysis is displayed on the bottom left panel.

Figure 13

Simulation results from Experiment 2 – Best model without the main mechanisms



Note. These curves were obtained by taking the best fitting model (see **Figure 12**, bottom left panel), and removing each mechanism one by one.

Discussion

Results from Experiment 2 show that the symmetrical serial position curves observed in Experiment 1 are likely due to the last item benefiting from strong recency. Once a distractor task is included between encoding and recall, serial position curves display a primacy effect across input position, regardless of the recall condition considered. These results have implications for studies using the forward/backward recall paradigm (Dougherty et al., 2023; Liu & Caplan, 2020), as they suggest that it might not be the best tool to disentangle the role of encoding- versus recall-related processes in WM. As the last item in the backward recall condition might be very strongly represented (for instance, by virtue of being in the focus of attention) at the time of recall, this could give the false impression that serial position curves are almost entirely explained by output interference. Using a distractor task as we did provides a quick fix to this problem.

The free-time benefit in the forward-backward recall paradigm is best explained by a model including a primacy gradient of activation, a consolidation mechanism, and a stabilization mechanism. The fact that the consolidation mechanism was favored in Experiment 2 contrasts with the results from Experiment 1. This issue is again likely due to the methodological limitations of Experiment 1. Despite the broad convergence of these results, one risk of our modeling approach is to end-up with a model which is task-specific. It is possible that while the combination of mechanisms selected through model comparison works for a forward-backward recall paradigm, it might fail at fitting results in a different paradigm. In the next experiment, we extend the free-time manipulation to a cued recall WM paradigm.

Experiment 3

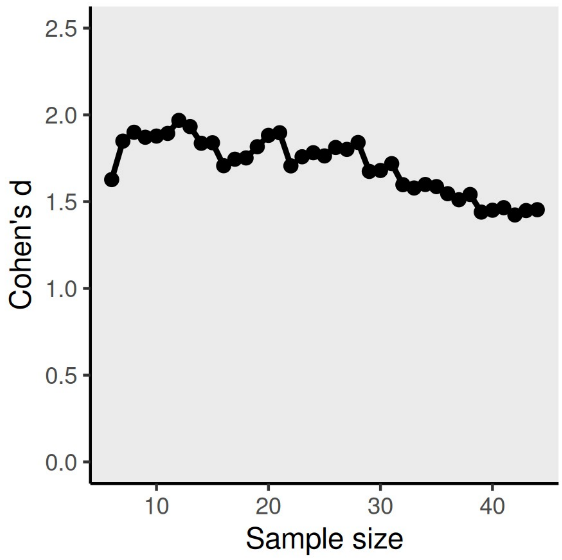
This experiment manipulates presentation rate in a cued recall paradigm. In cued recall, participants are presented with a list of items to be remembered and recall the items given a positional cue. Items are cued in random order. For instance, given the list “ABCDEF”, participants could be cued with “position 3”, and must recall “C”, followed by the cue “position 6” to which they have to answer “F”, and so forth. Contrary to the forward-backward recall paradigm used in the previous experiments, the cued recall paradigm has the advantage that items recalled first can be any item in the list, and not necessarily the first or last encoded items. This allows plotting memory performance as a function of input and output position in an independent manner, thereby providing a more direct picture of the free-time benefit for each serial position. Finally, Experiment 3 is a conceptual replication of the experiments conducted above. If the results observed in the previous experiments are robust, they should generalize to other paradigms.

Methods

Optional stopping rule. In Experiment 3, sample size was defined using the same optional stopping rule based on effect size stabilization as Experiment 1. Using this stopping rule, we reached a final sample size of 44 participants. The evolution of the effect size across sample size is displayed in **Figure 14**. As can be seen, the effect size stabilized again at d ~ 1.5.

Figure 14

Effect size stabilization – Experiment 3



Note. Cohen's ds were computed using the difference between the fastest (0.5 seconds / item) and slowest (4.0 seconds / item) time conditions.

Participants. Forty-four young adults aged between 18 and 35 years were recruited on the online platform Prolific (https://prolific.co/). All participants were English native speakers, reported no history of neurological disorder or learning difficulty, and gave their written informed consent before starting the experiment. The experiment had been approved by the ethics committee of the Faculty of Psychology at the University of Liège, project #2021-024.

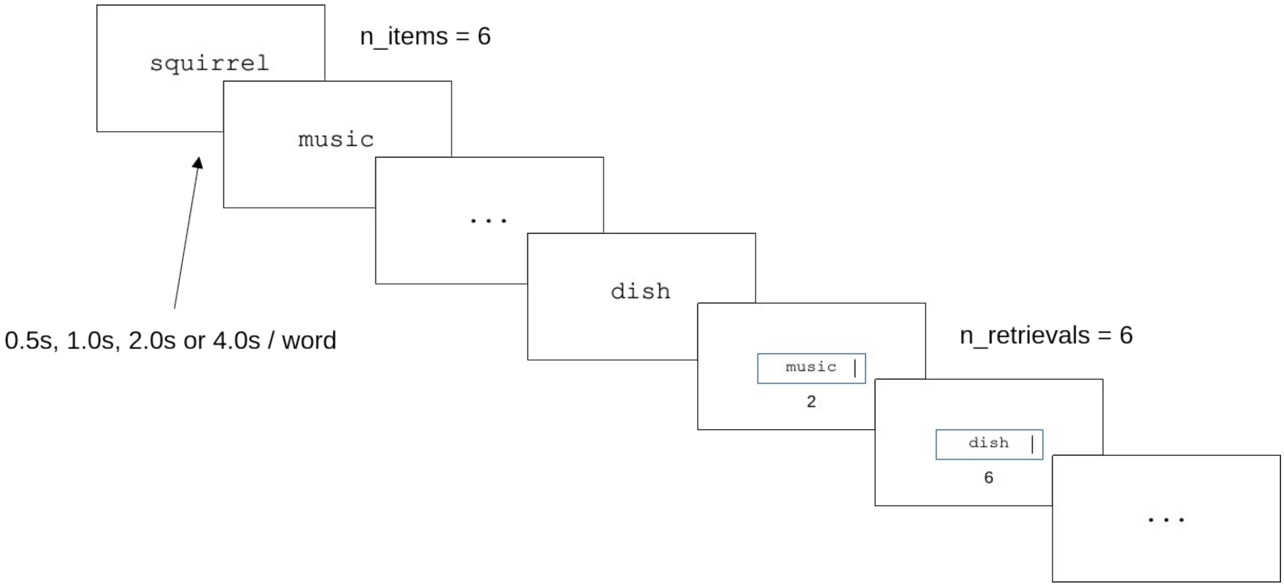
Design. Using the cued recall paradigm, we are left with four experimental conditions: four presentation rate conditions. All participants performed the four experimental conditions. Given the large number of possible recall combination (i.e., 6! = 720), results cannot be analyzed separately for each recall condition. Instead, results must be aggregated across recall conditions. Thanks to this aggregation process, memory performance can then be assessed across both input and output position.

Material. The lists were constructed using the same material and the same constraints as in Experiment 1 and Experiment 2. There were 16 lists per presentation rate condition, leading to a total of 64 experimental trials. Recall order was defined using the following constraints. For each presentation rate condition, we generated 8 recall orders by sampling from the numbers [1, 2, 3, 4, 5, 6] without replacement. We then checked that each input position was included at least once in each output position. If this criterion was reached, we repeated the process a second time.

Procedure. Participants were presented with six items to be remembered, with a varying presentation rate (i.e., 0.5, 1.0, 2.0 or 4.0 items per second). After the presentation of the last item to be remembered, participants were presented with a prompt box, along with a number below it indicating the position of the item they had to recall. For instance, if participants were presented with the items “tranquility, beer, square, uncle, stone, plague” (in this order), they could be presented with the cue “5”, to which they had to respond “stone”. They could then be presented with “1”, to which they had to respond “tranquility”, and so forth until all items were tested once. If participants could not remember an item, they were instructed to leave the prompt box empty. The rest of the procedure was identical to Experiments 1 and 2. The task is illustrated in **Figure 15**.

Figure 15

Setup used in Experiment 3



Note. Participants were visually presented with six items to be remembered. Presentation rate varied depending on the encoding time condition (i.e., 0.5, 1.0, 2.0, or 4.0 seconds per item). After the presentation of the last word to be remembered, participants were cued with a number below a prompt box, indicating the position of the word to-be-remembered. All items were tested once.

Scoring procedure. Participants' responses were scored as in Experiment 1.

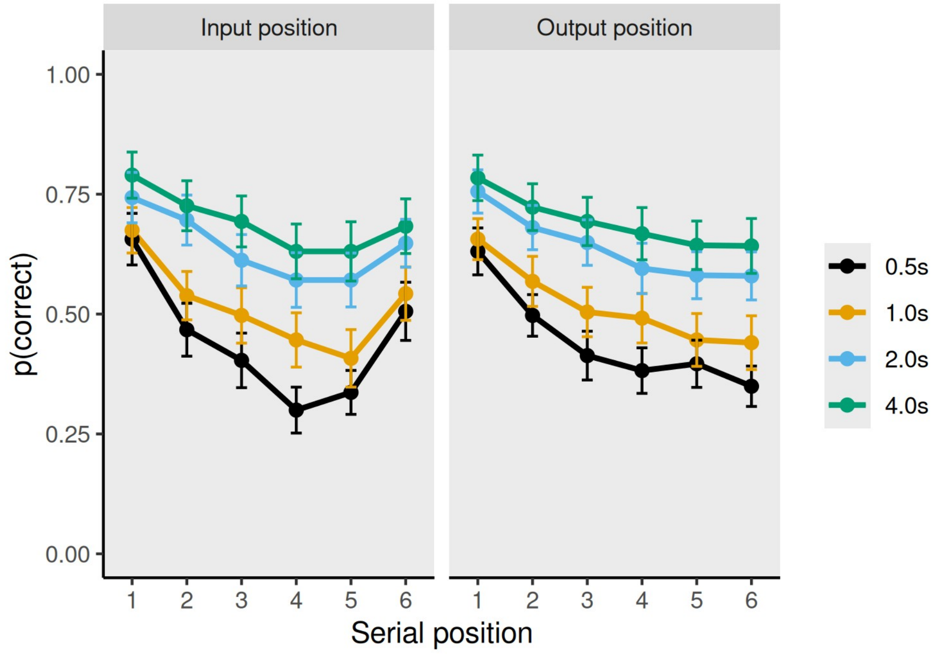
Statistical analyses. Statistical analyses were conducted using the same strategy as in Experiment 1.

Results

Recall performance as a function of input and output position is displayed in **Figure 16**. The Bayesian logistic regression model indicates that memory performance increased with slower presentation rates, and this was supported by decisive evidence when analyzed across input (BF_{10} = 4.939e+7) and output (BF_{10} = 4,695) position.

Figure 16

Empirical results – Experiment 3



Note. Error bars indicate 95% confidence intervals, corrected for within-subject variability.

Model fitting

We performed model comparison using the same top-down approach as used in Experiment 1 and Experiment 2. A summary of this analysis is reported in **Table 5**.

We started with a model including the consolidation, stabilization, and encoding-resource mechanisms. We compared this model against a model including the consolidation, stabilization, and primacy gradient mechanisms. Results indicate that the data were 358 times more likely under the model including the primacy gradient of activation than under the model including the encoding-resource mechanism. Predictions from these two models are reported in **Figure 17**. Again, the encoding-resource mechanism did not contribute much to the goodness-of-fit as compared to a primacy gradient of activation, suggesting a minor role in simulating the free-time benefit.

Next, we compared the model including the consolidation, stabilization and primacy gradient mechanisms against the same model without the primacy gradient mechanism. We found that the data were 469 times more likely under the model including the primacy gradient mechanism than under the model not including it. As in the other experiments, removing the primacy gradient mechanism resulted in no primacy effect across input position, as can be seen in **Figure 18**, under the "Without primacy" panel.

We next compared the model including the consolidation, stabilization and primacy gradient mechanisms against the same model without the stabilization mechanism. Results indicate that the data were 86 times more likely under the model including the stabilization mechanism than under the model not including it. Again, without this mechanism, the free-time benefit would be equally strong across serial position, leading to no fanning-out effect and therefore a wrong quantitative prediction, as can be seen in **Figure 18**, "Without stabilization" panel.

Finally, we compared the model including the consolidation, stabilization and primacy gradient mechanisms against the same model without the consolidation mechanism. The data were 50 times more likely under the model including the consolidation mechanism than under the model not including it. Consistent with Experiment 2, dropping the consolidation mechanism results in no free-time benefit for output position 1 (see **Figure 18**, "Without consolidation" panel), a result which is clearly observed when looking at **Figure 16**. Overall, the best-fitting model was the model including the consolidation, stabilization and primacy gradient mechanisms. Predictions from this model are illustrated in **Figure 17**, bottom panel.

Table 5. Model comparison from Experiment 3

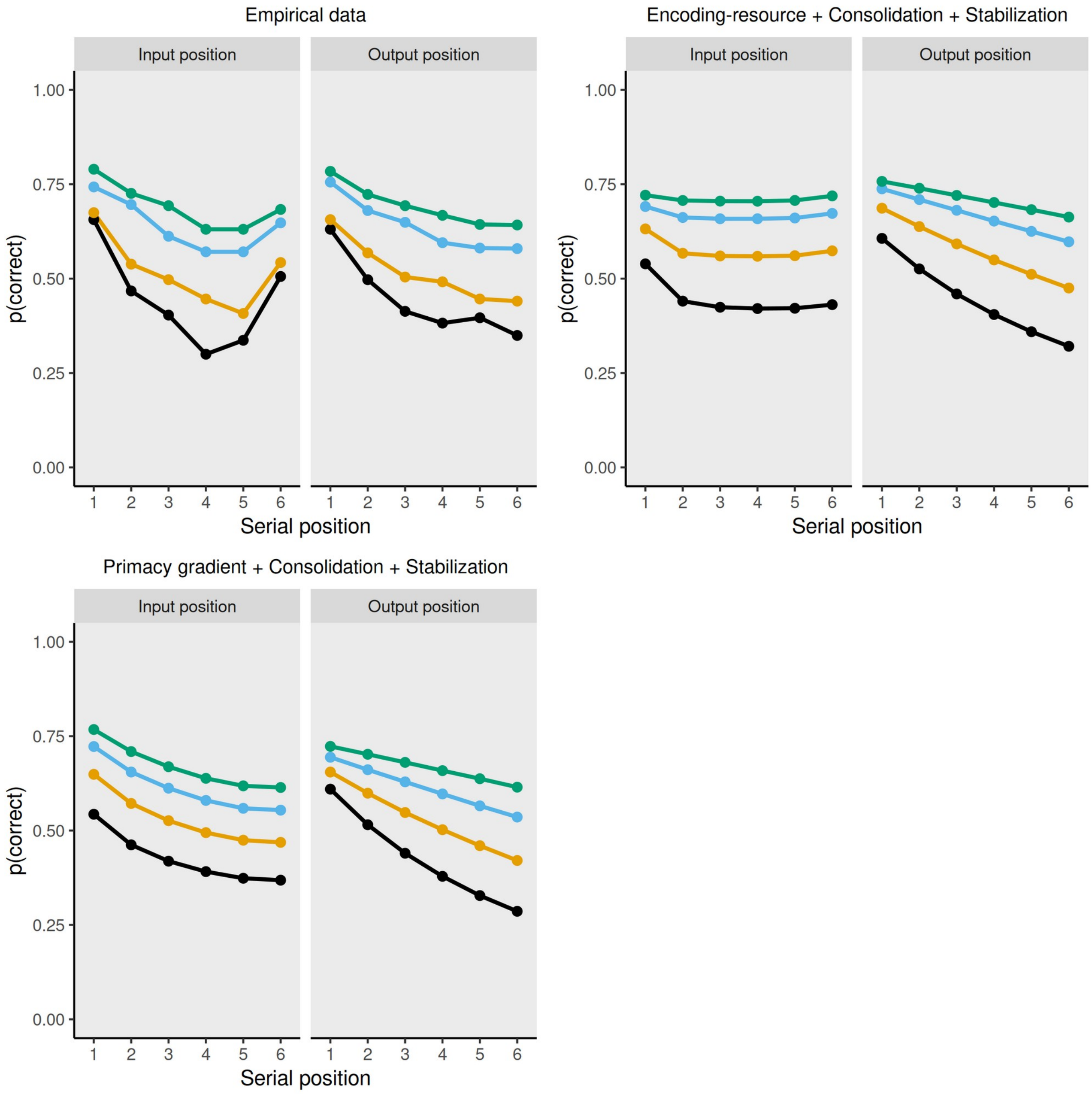
	Model 1	Model 2	$BIC_{m1} - BIC_{m2}$
Step 1	[consolidation] + [stabilization] + [encoding-resource]	[consolidation] + [stabilization] + [primacy gradient]	358
Step 2	[consolidation] + [stabilization] + [primacy gradient]	[consolidation] + [stabilization]	-469

Step 3	[consolidation] + [stabilization] + [primacy gradient]	[consolidation] + [primacy gradient]	-86
Step 4	[consolidation] + [stabilization] + [primacy gradient]	[stabilization] + [primacy gradient]	-50

Note. Model comparison was performed using a top-down approach, by considering first the most complex models justified by our experimental design and theories. Negative values indicate better fit for m1 as compared to m2.

Figure 17

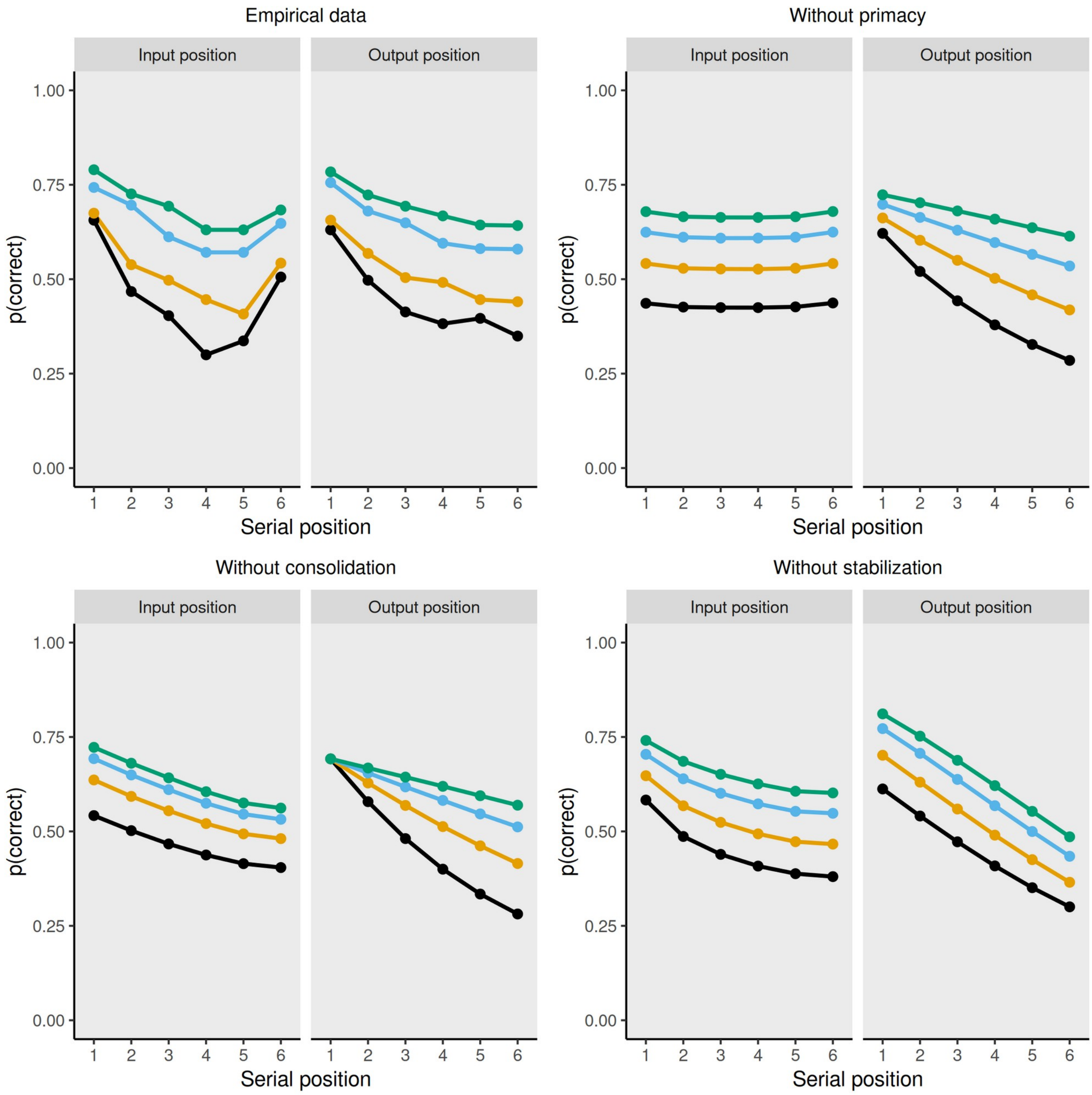
Model predictions from Experiment 3



Note. Upper left panel: Empirical data. Upper right panel: model including a consolidation, stabilization and encoding-resource mechanisms. Bottom left panel: model including a consolidation, stabilization and primacy gradient mechanisms. The winning model across the whole model comparison analysis is displayed on the bottom left panel.

Figure 18

Simulation results from Experiment 2 – Best model without the main mechanisms



Note. These curves were obtained by taking the best fitting model (see **Figure 17**, bottom left panel), and removing each mechanism one by one.

Discussion

Results of Experiment 3 converged with those observed in previous experiments: there was a clear free-time benefit. Model comparison showed that a model including the consolidation and stabilization mechanisms was necessary to account for the data. Similarly, we did not find evidence that an encoding-resource mechanism was necessary to account for the free-time benefit. Instead, a primacy gradient of activation mechanism was critical to produce a primacy effect across input position. In the next section, we explore the ability of another model to account for the free-time benefit, namely the TBRS* model.

Additional Simulations: TBRS*

In this section, we focus on the ability of decay and refreshing models to capture the free-time benefit using the TBRS* architecture, a connectionist implementation of the Time-Based Resource Sharing Theory (TBRS, Barrouillet et al., 2004). We chose to use an existing connectionist model to test the decay and refreshing account of the free-time benefit, because the dynamics associated with these maintenance-related processes are too complex to be captured by a closed-form expression. Oberauer (2022) provided a simple mathematical formalization of what should theoretically happen in a decay and refreshing model, but he did not directly simulate the dynamics of these mechanisms.

In the TBRS theory, items constantly decay when out of attention. Decay can be counteracted by going back to a previously encoded item using the focus of attention, a central bottleneck limited to one item. The TBRS* model is the most advanced description of the TBRS theory, as it makes explicit assumptions regarding every mechanism implemented in its core architecture. The model describes how (i.e., via which processes) items are refreshed, for how long, and in which order, a feature absent in other verbal or mathematical descriptions of the theory (e.g., Barrouillet et al., 2004; Gauvrit & Mathy, 2018). As TBRS is a theory in which time plays an important role, it is the best candidate to model the free-time benefit when considering a decay and refreshing perspective. Although the TBRS* model has been used in the context of many WM paradigms (Kowialiewski et al., 2021; Kowialiewski, Lemaire, et al., 2024; Lemaire et al., 2021; Lemaire & Portrat, 2018; Portrat et al., 2016; Portrat & Lemaire, 2015), whether it can simulate the free-time benefit in immediate serial recall has not yet been tested.

We will not re-describe the mathematical implementation of TBRS* in this manuscript. The reader interested in the exact implementation can read the original publication in Oberauer & Lewandowsky (2011). The publications by Lemaire and Portrat (2018) and Portrat et al. (2016) also contain useful information and illustrations. We provide here a verbal description of the general principles of TBRS* to keep the manuscript accessible enough.

Model description

As in many WM models (including those presented above), encoding in TBRS* is done by binding items to positional contexts. Bindings are formed by creating new associations using rapid Hebbian learning. These associations are stored in a weight matrix. When out of attention, the item-context associations in the weight matrix continually decay by scaling them by a constant proportion which depends on the elapsed time. If there is free-time available between successive encoding periods, the deleterious effect of decay can be counteracted by restoring the item-context associations. This is done using refreshing.

Refreshing starts by retrieving the desired item after cueing it with its positional context. Hence, the same mechanisms are involved in both refreshing *and* recall. This implies that refreshing is subject to potential failure, such as retrieving a wrong list-item (i.e., transposition error) or the inability to retrieve an item at all (i.e., omission error). This property is precisely the reason why the dynamics of the model are complex and must be simulated (i.e., a closed-form expression is difficult to derive). After identifying the item to be refreshed, it is re-encoded to the context which served as a cue to retrieve it using the same rapid Hebbian learning used during encoding. In TBRS*, the time dedicated to refreshing an item is set at 80ms (Vergauwe & Cowan, 2015)[5]. During the time spent refreshing an item, all the other representations decay, because they are out of attention. After the encoding phase, the model recalls items by retrieving them one by one using the same mechanism as those used during refreshing. When recalling an item, all WM representations decay due to the mere passage of time. This encoding duration is fixed to 500 ms by default.

In TBRS*, four sources of errors are possible. First, the model can recall a non-target item that was part of the list (i.e., transposition error). This phenomenon is due to both the positional uncertainty and the noise added at retrieval. Second, there is some probability to recall an item which was not part of the list, resulting in an extra-list intrusion. This occurs because when cueing an item from its context, this generates a pattern of activation to which some noise is added. The noise is also added to items that were not part of the list, resulting in a non-null probability that the activation level of a non-list item wins the competition. Third, the model can fail to retrieve an item, resulting in an omission error. Omissions are modeled by determining an omission threshold. At retrieval, if all item activations are below the omission threshold, there will be no output. Fourth, the model can recall an item which has already been recalled, resulting in a repetition error. To prevent repetitions, the model implements a form of response suppression called *removal,* which is done by performing Hebbian anti-learning. Hebbian anti-learning is equivalent to encoding via rapid Hebbian learning, except that a negative learning rate is used, thus removing the item from WM.

[5]Note that a recent empirical exploration of refreshing estimates it to 200 ms (Oberauer & Souza, 2020). In these simulations, we stick with the 80 ms value to stay as close as possible to the original implementation.

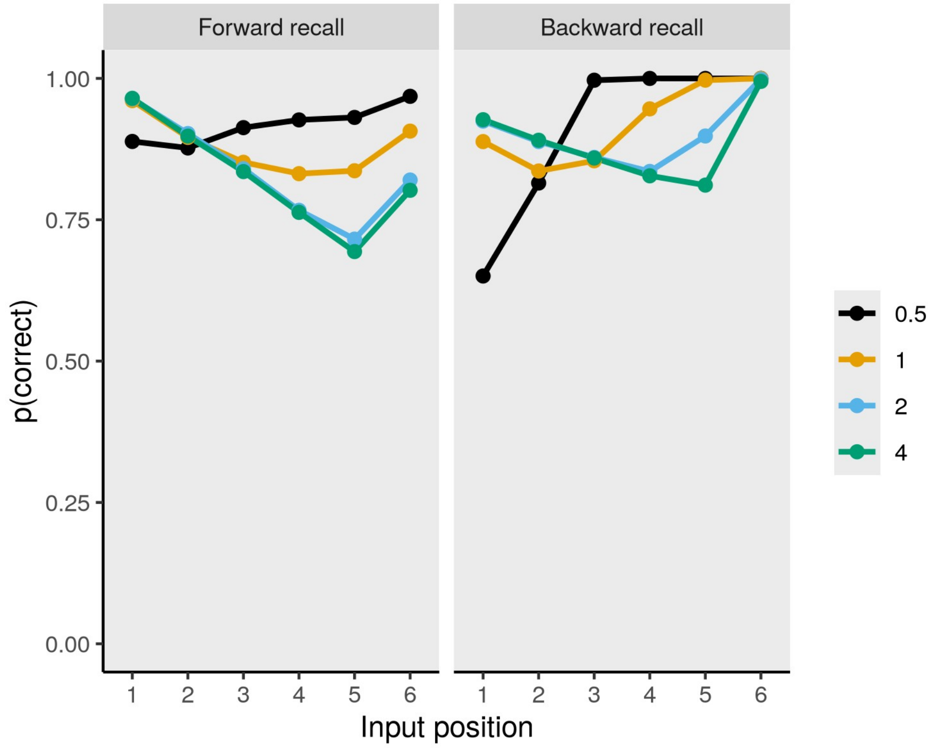
There are different ways the model can refresh items. Different schedules have been proposed, such as refreshing items cumulatively (1, 2 – 1, 2, 3 – 1, 2, 3, 4…) (Vergauwe et al., 2016), or refreshing the least activated item (Lemaire et al., 2018). In these simulations, we set the refreshing schedule to cumulative as implemented in the original TBRS* architecture. Basically, the model starts by cueing the item with the first positional cue, re-encodes the item, and then performs the same operation using the next positional cue (i.e., position 2), and so forth until the last encoded item. The model then cycles back to the first item and continues this loop until the next to-be-remembered item appears. This way of refreshing items creates a primacy gradient of activation without the need for an explicit mechanism (Oberauer & Lewandowsky, 2011).

Results

Figure 19 reports simulations results from TBRS*. The model was run using the default parameter values (see **Appendix B**). The TBRS* model predicts a free-time disadvantage: Slower presentation rate leads to poorer recall performance. This pattern of results may appear surprising at first. We tested whether this property of the TBRS* model is specific to the set of parameters used in the current simulations, or due to a general property of the model. We ran a grid search covering a broad range of the parameter space, involving 16,384 data points. Each data point involved 2,500 simulated trials. The model was therefore run for a total of 40,960,000 trials. The range of values and associated parameters are reported in **Table B2** (see **Appendix B**). In this grid search, the model performed immediate serial recall in forward order across all four time conditions. Results from this grid search indicate that the model got better with slower presentation rates in only 2.9% of the explored parameter space, and got worse in 66.6%. The remaining percentages represents cases where the direction of the effect was not systematic across the four conditions. We therefore reach the conclusion that TBRS* predicts a free-time disadvantage, which constitutes a general property of the model.

Figure 19

Simulation results from the TBRS model*

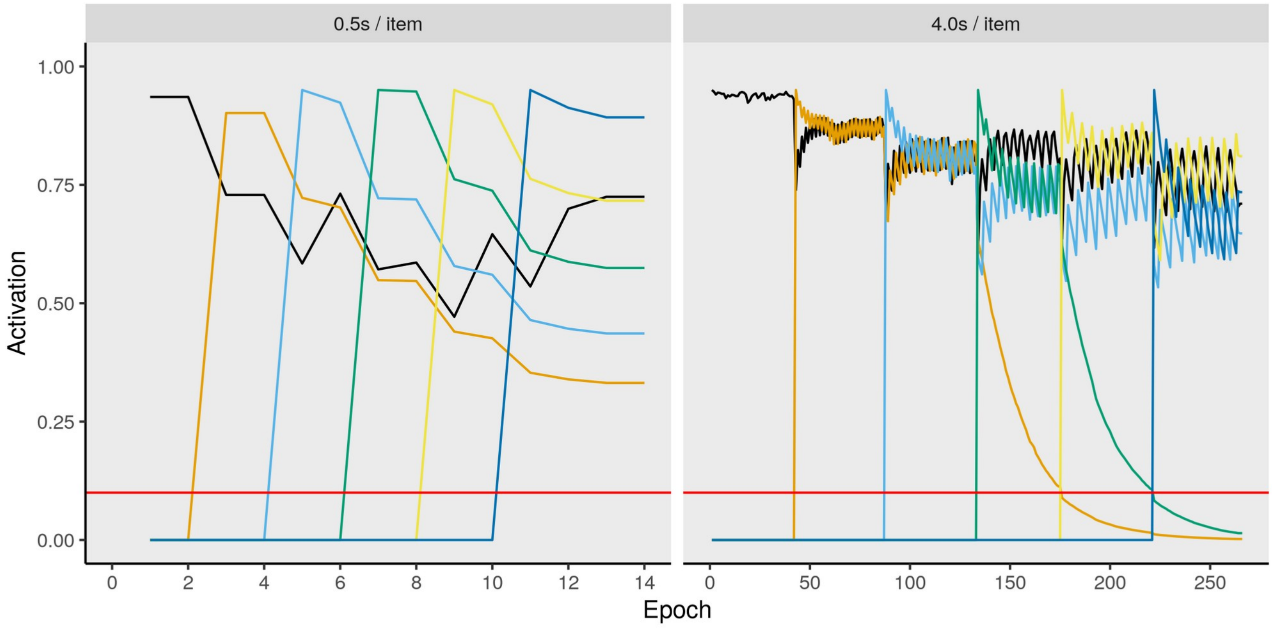


Note. The model was run using the default parameter values.

Why does TBRS* predict a free-time disadvantage? First, faster presentation rates lead to less time-based forgetting. Consider the extreme case in which items are presented as quickly as 0.5s / item, such as in the present study. In this configuration, three seconds occur between the beginning of the presentation of the first item and the retrieval phase. These three seconds are not enough to cause strong time-based forgetting. Hence, items are still very active right before recalling them. Second, more free-time between two encoding periods induces forgetting due to refreshing. Because in TBRS* the process of refreshing an item is equivalent to retrieving it, increased refreshing opportunities also means increased occurrence of retrieval-based errors. This is an unintuitive consequence of refreshing which is often overlooked: Refreshing is supposed to counteract the deleterious effect of decay. But because refreshing (and therefore retrieval) is necessarily error-prone, more refreshing also results in a higher absolute number of errors, a property already identified by Lewandowsky & Oberauer (2015). To illustrate this, we display in **Figure 20** the time-course of items' activation over one trial in fast and slow presentation rate conditions. In the fast encoding-time condition, no item has been forgotten over the whole trial, and all of them are available at retrieval, as indicated by their activation level well-above the omission threshold (i.e., the red horizontal line). The only possible errors are therefore those occurring during recall. In contrast, in the slow encoding-time condition, although items are maintained through refreshing, some of them are dropped from the competition due to retrieval-related errors. In **Figure 20**, right panel, items 2 and 5 are dropped from the competition, which is indicated by their activation value decaying towards zero. During the recall phase, the model is therefore left with only 4 items. This is in contrast with the fast presentation rate condition in which all items are still available at retrieval. This mere pattern leads to better recall performance in the fast encoding time condition.

Figure 20

Time-course of the model over one trial

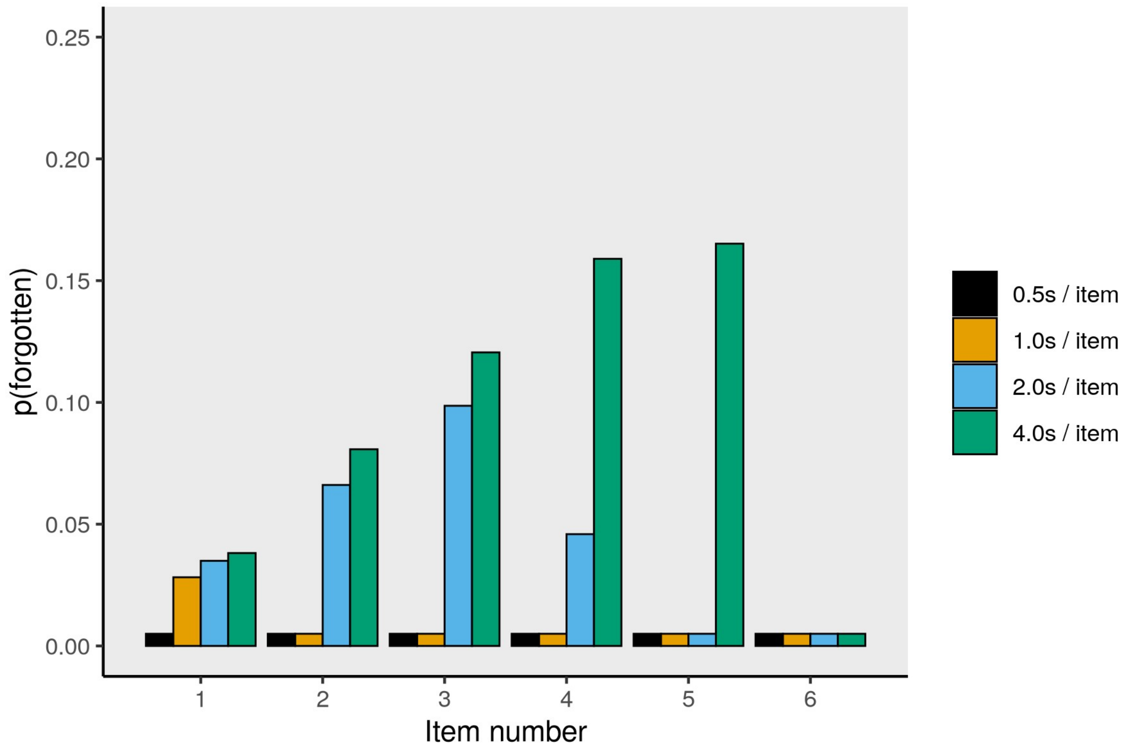


Note. Left panel: fast presentation rate condition. Right panel: slow presentation rate condition. Activation values were extracted by averaging values of weights connecting items to their contexts. The x-axis represents the discrete simulated steps in the model, although all processes are implemented in a continuous way in the mathematical equations.

To demonstrate the argument exposed in the previous paragraph, **Figure 21** displays the proportion of forgotten items (i.e., items below the omission threshold) after the end of the encoding phase (i.e., right before the model starts recalling the items) for each item (1 through 6) and across presentations rates (0.5, 1.0, 2.0 and 4.0s / item). These proportions result from 100,000 simulated trials using the standard parameters of the model. As illustrated, the proportion of items forgotten during encoding/maintenance was virtually zero in the fastest condition. As presentation rate slowed down, this proportion increased for all items. Thus, additional free-time does not necessarily benefit WM performance in decay and refreshing models.

Figure 21

Proportion of items forgotten right before recalling the items in the TBRS model*

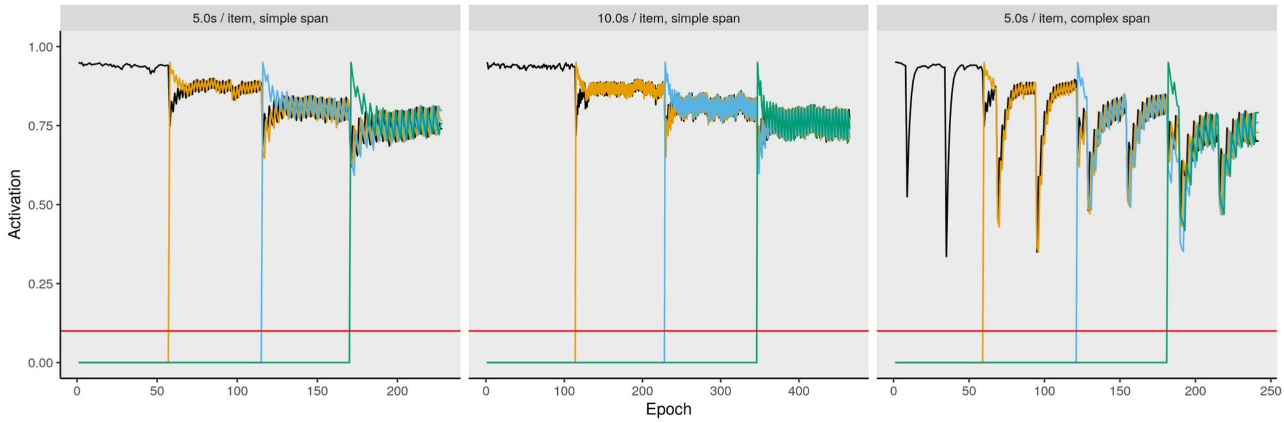


Note. Proportions were computed over 100,000 simulated trials for each encoding time condition. An item was considered forgotten if its activation level fell below the value corresponding to the omission threshold.

The fact that we failed to reproduce the free-time benefit in immediate serial recall using TBRS* seems to be at odd with multiple studies which successfully modeled the cognitive load effect with it (Lemaire et al., 2018, 2021; Lemaire & Portrat, 2018; Oberauer & Lewandowsky, 2011; Portrat & Lemaire, 2015). Basically, when people process distractors between memoranda, they forget more items, and this forgetting increases with the time they spend processing distractors. We also reproduce the cognitive load effect in TBRS*, as reported in **Appendix C**. Why does this contradiction occur? Cognitive load is usually manipulated in designs in which a long period of time occurs between the occurrence of the first item and the recall phase. In such slow-paced designs, additional free-time doesn’t affect TBRS*, as shown in **Figure 22**, left and middle panels. This is because with long retention intervals, refreshing perfectly compensates decay, and the model reaches a stable equilibrium. Introducing distractors breaks this equilibrium, causing additional forgetting, as illustrated in **Figure 22**, right panel. This phenomenon drives the cognitive load effect. In immediate serial recall however, the very fast presentation rate of the items prevents them from decaying, reducing forgetting. Thus, presenting items more rapidly in the absence of distractors improves memory performance in the model, which explains this apparent contradiction.

Figure 22

Time course in the model



Note. Comparing the left and middle panels, adding more free-time in a slow-paced experiment design doesn’t improve memory performance, because decay and refreshing perfectly balance each other, leading to an equilibrium. Adding two distractors per item (right panel) disrupts this equilibrium.

To sum up, these additional simulations show that TBRS* cannot simulate the free-time benefit in immediate serial recall. This occurs because faster presentation rates compensate for the deleterious effect of decay, and more free-time causes more opportunities to produce retrieval-based errors during maintenance via refreshing (see also Lewandowsky & Oberauer, 2015).

General Discussion

This study aimed to understand how giving people more time to study items affects the way they memorize and subsequently recall them serially. To achieve this, we used a combined method involving behavioral experiments and computational modeling for providing a comprehensive assessment of the plausibility of the different candidate-mechanisms that have been forwarded to explain the free-time benefit. Our results consistently favor two mechanisms. The first one is a consolidation mechanism based on the re-encoding of the just-encoded item. The second one is a stabilization mechanism whereby items resist more strongly to output interference with additional free time.

Where Does the Free-Time Benefit Come From?

In contrast to our results, the original simulation work reported by Oberauer (2022) did not favor a consolidation mechanism in which the just-encoded item continues to be encoded. This discrepancy stems from the fact that the forward-recall paradigm used by Oberauer was not optimal to show the manifestation of such a mechanism. In our data, this is demonstrated by a systematic recall advantage for items encoded first. In addition to this consolidation mechanism, our results support a mechanism in which additional free time benefits WM performance by stabilizing the just-encoded item, which becomes more robust to output interference. It is conceivable that both the consolidation and stabilization mechanisms occur simultaneously when people re-encode the item they just saw, or that one mechanism is the consequence of the other. One limitation of our approach is that these mechanisms lack a more precise implementation. This choice is a consequence of our strategy to use a standard model of WM in which only the commonly accepted assumptions of most WM models were incorporated (Cowan et al., 2002; Henson, 1998; Hurlstone & Hitch, 2015; Lewandowsky, 1999; Page & Norris, 1998). In order to fully test the plausibility of these mechanisms, they should be implemented in more detailed architectures in the future.

Candidates architecture for this endeavor involve models postulating that items are encoded in WM using feature vectors, such as interference models (Oberauer et al., 2012; Oberauer & Lin, 2024), or more recently the Revised Feature Model (Saint-Aubin et al., 2021). In these models, items are encoded as feature vectors by associating them to positional markers. Items are retrieved by first cueing them with their position, as classically done in most WM models. However, because these features are subject to interference to a varying degree, they cannot be recalled as such, but must be compared to items stored in long-term memory (i.e., in the lexicon). During this comparison process, the best-matching vectors are most likely to be selected for output. One way these models could account for the free-time benefit is by adding the assumption that people refocus their attention on the just-encoded item, thus partially restoring the vector to its initial configuration or by re-encoding it more strongly. One additional consequence could also be to enrich the vector with additional features that aren't initially encoded (McClelland & Chappell, 1998; Ricker & Vergauwe, 2022; Shiffrin & Steyvers, 1997), possibly through a deeper semantic encoding. This could especially be the case for those items that would benefit the most from a deeper encoding process, such as concrete words. This hypothesis would align with Oberauer (2022)'s observation that concrete words benefit more from additional free time than abstract words do. Hence, our findings pave the way for future modeling work, especially the way these mechanisms could be integrated in broader architectures.

It must be noted that the Revised Feature Model currently includes a rehearsal mechanism operating during the free-time available between memoranda. A recent study suggests that such a rehearsal mechanism can explain patterns of results observed in the production effect (Dauphinee et al., 2024). However, this rehearsal mechanism, by itself, cannot simulate the free-time benefit because people still benefit from additional free time under concurrent articulation (Oberauer, 2022).

It is important to note that our method does not allow for exploring differences regarding the best-fitting model at an individual level. It is possible that the type of best-fitting model(s) varies across participants. Answering to that question would however require increasing the number of trials per participants so as to minimize intra-individual variability. Therefore, our approach is currently limited to tell whether a particular mechanism is being supported or not, generally speaking. Based on fine-grained modeling works, it would be important in future studies to establish to what extent participants differ in their encoding strategies, as recently shown (Bartsch et al., 2024).

The Encoding-Resource Mechanism

Although our series of experiments did not favor an encoding-resource mechanism, a recent study suggests its existence. Recently, Mizrak and Oberauer (2021) proposed a gap manipulation paradigm in which participants encoded lists of letters and recalled them serially. The gap manipulation involved the inclusion of a pause between two items, which occurred at an unexpected inter-item list position. To control for the overall impact of temporal grouping (Ryan, 1969), Mizrak and Oberauer included two gap conditions, one involving a short gap and another one involving a long gap. Comparison between both gap conditions showed that the long gap manipulation improved memory performance for the items *following* the gap, creating a proactive benefit. Memory performance slightly increased only for the item directly preceding the longer gap, producing a limited retroactive benefit. These results support the encoding-resource mechanism, as it is the only one predicting a proactive benefit, but no retroactive benefit. At the same time, the results of Mizrak and Oberauer conflict with other observations. Maybery et al. (2002) performed a similar experiment involving long and short gaps, except that in the long gap condition, items were presented more rapidly within each group. Contrary to the encoding-resource mechanism's predictions, they observed that the long gap manipulation increased memory performance globally, producing proactive and retroactive benefits. If anything, an encoding-resource mechanism should have predicted worse memory performance for items at the beginning of the list in the long gap condition due to the lack of resource replenishment with faster presentation rates. Similarly, Ryan (1969) failed to show that increasing the gap in temporal grouping manipulations improved memory performance. One feature of these studies is the fact that they all involve a form of temporal grouping. This is a limitation, because current models of WM fall short at providing a fully satisfying explanation of this phenomenon (see Gorin, 2021), which also means that potential interactions between temporal grouping and additional free time are currently outside of our understanding. One way to resolve these current contradictions is to better understand what mechanisms drive the temporal grouping effect, coupled with a deeper behavioral exploration of the gap manipulation. Overall, our study does not rule out the existence of an encoding-resource mechanism. It mainly shows that this mechanism does not better account for the data we report here. There are also variations of this mechanism we did not consider. For instance, Ricker and Vergauwe (2022) recently suggested the possibility that the encoding-resource mechanism could limit the amount of enrichment of a given memory representation. Hence, our results suggest that if the encoding-resource mechanism exists, it is not the only mechanism accounting for the free-time benefit. This conclusion contrasts with the one Oberauer (2022) reached, for whom the encoding-resource mechanism was the most plausible among all the tested alternatives.

The Question of Decay and Refreshing

Simulations reported by Oberauer (2022) showed that a decay and refreshing model cannot account for the fanning-out effect, as these models produced a reversed fanning-out effect (i.e., bigger free-time benefit for early than late items). The present study goes even further by showing that the TBRS* model implementing decay and refreshing processes in a more realistic manner predicts a *reversed* free-time benefit. This is explained by the fact that faster presentation rates imply less time-based decay, thus producing an opposite effect. This result contradicts the intuitive idea that providing more free-time should necessarily increase refreshing opportunities, which would in turn counteract the deleterious effect of decay. The free-time benefit is not the only time-based phenomenon that TBRS* fails to simulate. Farrell et al. (2016) performed a series of experiments involving complex span tasks. Farrell and colleagues manipulated the frequency of distractors appearing at specific serial positions and observed that this manipulation affected memory performance locally. In contrast to this, distractors affect items globally in TBRS*, because decay affects all items simultaneously when attention is driven away from memoranda. In addition, a recent study found an absence of cognitive load effect in a Brown-Peterson paradigm (Langerock et al., 2024), a result which is also difficult to reconcile with the TBRS theory whose foundations lie in the cognitive load effect. Together with our results, these studies present challenges for decay and refreshing models. Addressing these issues is important to improve this family of models in the future.

The way forward is to revise some of the assumptions implemented in TBRS*, and test if such revisions have any improvement on the model's predictions. For instance, one could consider the possibility that items decay less rapidly after each refreshing opportunity. With this assumption, additional free time would result in fewer decay and therefore better recall performance in slow as compared to fast presentation rates. It is important to not completely reject a whole family of models based on the fact that a specific implementation fails at accounting for some benchmarks. Despite their limitations, decay and refreshing models offer a sound explanation for other important phenomena. First, they naturally produce a primacy gradient of activation without the need for a specific mechanism. Second, they do not require an explicit output interference mechanism either, because the time spent recalling the items already produces this effect. Third, they explain why the last-encoded item is very strongly represented into WM, a phenomenon also observed in our simulations (see **Figure 19**).

Conclusion

Recent studies suggest that presenting items at a slower presentation rate improves memory performance only in a proactive manner. In contrast to this, our experiments involving a broader range of experimental conditions reveal that slower presentation rates improves memory performance across the entire list. Our computational modeling work indicates that the two most plausible explanations for this free-time benefit involve the re-encoding of the just-presented item and a stabilization mechanism that mitigates the impact of output interference. While challenging existing accounts of the free-time benefit, our results offer promising prospects for future developments of computational accounts of WM and the free-time benefit.

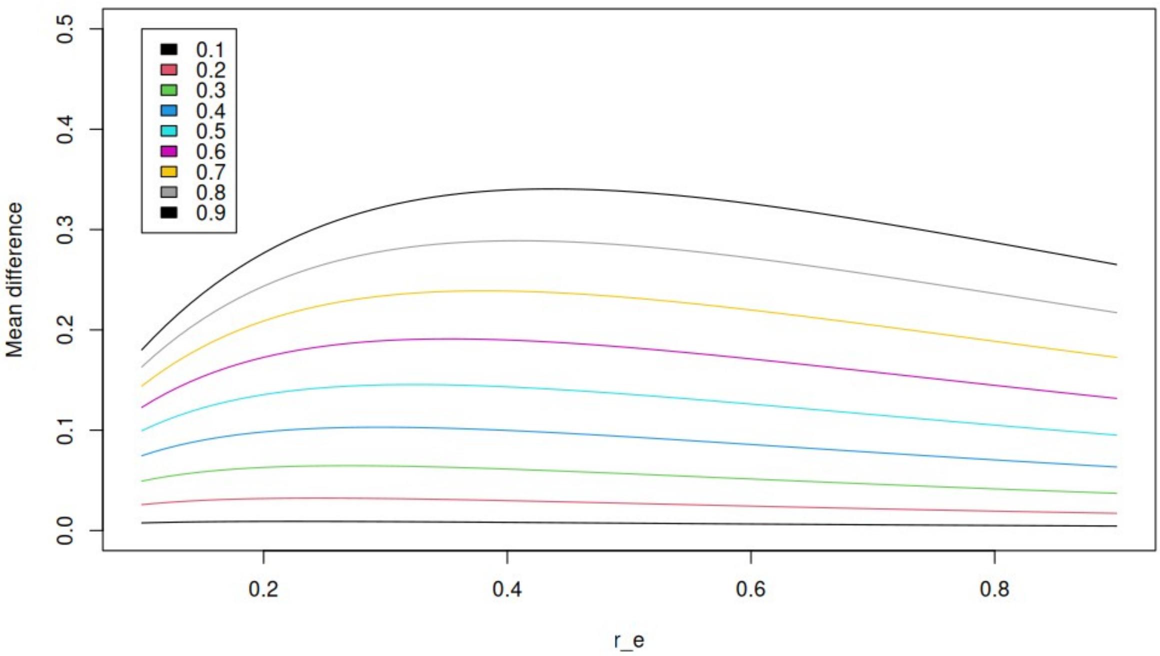
CRediT Statement:

BK: Conceptualization; Data curation; Formal analysis; Funding acquisition; Project administration; Investigation; Methodology; Software; Validation; Visualization; Roles/Writing - original draft; and Writing - review & editing

SM: Conceptualization; Resources; Supervision; Funding acquisition; Methodology; Validation; Roles/Writing - original draft; and Writing - review & editing

Appendix A

To define the r_e value of 0.44 used for the encoding-resource mechanism, we performed a grid search over the p_r and r_e parameters, comparing a fast (1 second/item) and slow (4 second/item) presentation rate. We computed the mean difference between these two presentation rate conditions for each point in the parameter space. Both parameters were bounded between 0.1 and 0.9. Results of this grid search are displayed in the figure below:



The Y-axis represents the mean difference between the fast and slow conditions: the higher the value, the bigger the free-time benefit. Each line represents a different p_r value, as indicated by the legend. Finally, each point on the X-axis represents a different r_e value. This figure indicates that fixing the r_e parameter to 0.44 and estimating p_r freely allows the mechanism to cover the broadest range of possible outcomes. This r_e value was chosen by looking at the top line, which corresponds to a p_r value of 0.9. For this parameter value, the maximum possible observable free-time benefit is located at r_e = 0.44. With this configuration, the model can produce a mean encoding strength difference ranging from nearly 0.0 (if p_r is set to 0.1) to around 0.35 (if p_r is set to 0.9).

Appendix B

Table B1. Default values used in the TBRS* simulation

Parameter	Meaning	Value
R	Processing rate	6.0
σ	Noise added at retrieval	0.02
θ	Omission threshold	0.1
D	Decay rate	0.5
P	Position marker overlap	0.3
s	SD of processing rate	1.0
T_r	Refreshing duration	80 ms

Note. Values were taken from the original Oberauer and Lewandowsky (2011) study.

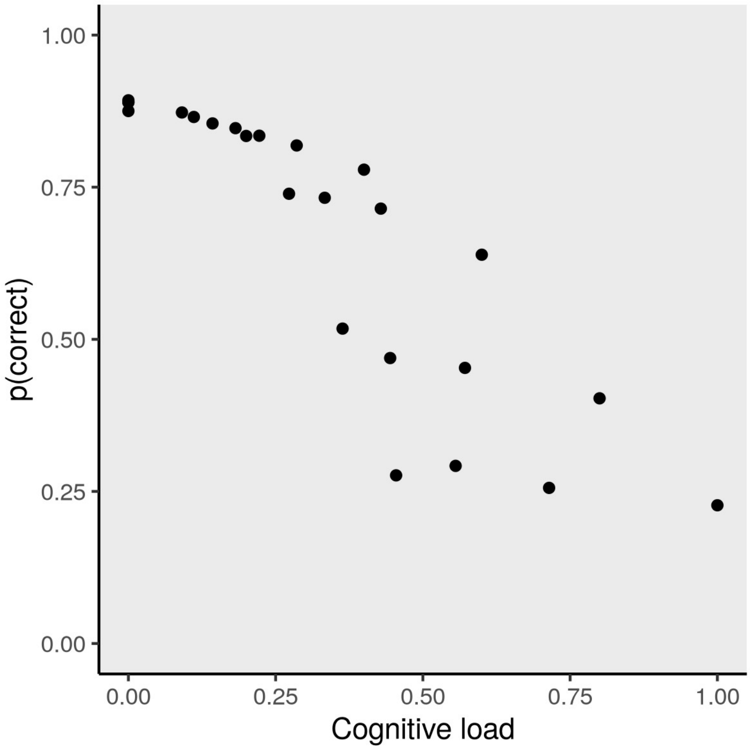
Table B2. Range of values used in the grid search

Parameter	Meaning	Range
R	Processing rate	[1.0 – 6.0] step = 0.714
σ	Noise added at retrieval	[0.0 – 0.5] step = 0.071
θ	Omission threshold	[0.0 – 0.6] step = 0.086
D	Decay rate	[0.1 – 0.9] step = 0.114

Note. Steps were chosen to generate eight different values for each parameter.

Appendix C

Cognitive load effect in TBRS*



Note. The model was launched using 5 items and presentation rates of [3.0, 4.0, 5.0, 6.0]. During this total time, attention was occupied 0.5 seconds dedicated to encoding, and the model always processed three distractors for periods corresponding to [0.0, 0.25, 0.5, 0.75, 1.0, 1.25]. The two variables (presentation rate and time to process distractors) were manipulated orthogonally, leading to a total 24 cognitive load levels. Each cognitive load level was run using 10,000 simulations.

References

Anderson, R. B., Crawford, J. C., & Bailey, M. H. (2022). Biasing the input: A yoked-scientist demonstration of the distorting effects of optional stopping on Bayesian inference. *Behavior Research Methods*, *54*(3), 1131–1147. https://doi.org/10.3758/s13428-021-01618-1

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

Bartsch, L. M., Souza, A. S., & Oberauer, K. (2024). The benefits of memory control processes in working memory: Comparing effects of self-reported and instructed strategy use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001370

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576. https://doi.org/10.1037/0033-295X.114.3.539

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Cowan, N. (1999). An Embedded-Processes Model of working memory. In *Models of working memory: Mechanisms of active maintenance and executive control* (Cambridge University Press, pp. 62–101).

Cowan, N., & Hardman, K. O. (2021). Immediate recall of serial numbers with or without multiple item repetitions. *Memory*, 1–18. https://doi.org/10.1080/09658211.2021.1942920

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding Serial Recall. *Journal of Memory and Language*, *46*(1), 153–177. https://doi.org/10.1006/jmla.2001.2805

Dauphinee, I., Roy, M., Guitard, D., Yearsley, J. M., Poirier, M., & Saint-Aubin, J. (2024). Give me enough time to rehearse: Presentation rate modulates the production effect. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02437-5

Dougherty, M. R., Halpern, D., & Kahana, M. J. (2023). Forward and backward recall dynamics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(11), 1752–1772.

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*(2), 223–271. https://doi.org/10.1037/a0027371

Farrell, S., Oberauer, K., Greaves, M., Pasiecznik, K., Lewandowsky, S., & Jarrold, C. (2016). A test of interference versus decay in working memory: Varying distraction within lists in a complex span task. *Journal of Memory and Language*, *90*, 66–87. https://doi.org/10.1016/j.jml.2016.03.010

Gauvrit, N., & Mathy, F. (2018). Mathematical transcription of the 'time-based resource sharing' theory of working memory. *British Journal of Mathematical and Statistical Psychology*, *71*(1), 146–166. https://doi.org/10.1111/bmsp.12112

Gorin, S. (2021). EXPRESS: Temporal grouping effects in verbal and musical short-term memory: Is serial order representation domain-general? *Quarterly Journal of Experimental Psychology*, 174702182110574. https://doi.org/10.1177/17470218211057466

Guérard, K., & Saint-Aubin, J. (2012). Assessing the effect of lexical variables in backward recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 312–324. https://doi.org/10.1037/a0025481

Guérard, K., Saint-Aubin, J., Burns, S. C., & Chamberland, C. (2012). Revisiting backward recall and benchmark memory effects: A reply to Bireta et al. (2010). *Memory & Cognition*, *40*(3), 388–407. https://doi.org/10.3758/s13421-011-0156-2

Henson, R. N. A. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive Psychology*, *36*(2), 73–137. https://doi.org/10.1006/cogp.1998.0685

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal Grouping Effects in Immediate Recall: A Working Memory Analysis. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 116–139. https://doi.org/10.1080/713755609

Hurlstone, M. J., & Hitch, G. J. (2015). How is the serial order of a spatial sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 295–324. https://doi.org/10.1037/a0038223

Jolicœur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Kowialiewski, B., Krasnoff, J., Mizrak, E., & Oberauer, K. (2023). Verbal working memory encodes phonological and semantic information differently. *Cognition*, *233*(105364). https://doi.org/10.1016/j.cognition.2022.105364

Kowialiewski, B., Lemaire, B., & Portrat, S. (2021). How does semantic knowledge impact working memory maintenance? Computational and behavioral investigations. *Journal of Memory and Language*, *117*(104208). https://doi.org/10.1016/j.jml.2020.104208

Kowialiewski, B., Lemaire, B., & Portrat, S. (2024). Similarity-Based Compression in Working Memory: Implications for Decay and Refreshing Models. *Computational Brain & Behavior*, *7*, 163–180. https://doi.org/10.1007/s42113-023-00179-0

Kowialiewski, B., Majerus, S., & Oberauer, K. (2023). Does semantic similarity affect immediate memory for order? Usually not, but sometimes it does. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *50*(1), 68–88. https://doi.org/10.1037/xlm0001279

Kowialiewski, B., & Oberauer, K. (2024). Testing the response suppression mechanism of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001359

Langerock, N., Oberauer, K., Throm, E., & Vergauwe, E. (2024). *The cognitive load effect in working memory: Refreshing the empirical landscape, removing outdated explanations*. https://doi.org/10.31234/osf.io/3zbe2

Lemaire, B., Heuer, C., & Portrat, S. (2021). Modeling Articulatory Rehearsal in an Attention-Based Model of Working Memory. *Cognitive Computation*, *13*(1), 49–68. https://doi.org/10.1007/s12559-020-09791-9

Lemaire, B., Pageot, A., Plancher, G., & Portrat, S. (2018). What is the time course of working memory attentional refreshing? *Psychonomic Bulletin & Review*, *25*(1), 370–385. https://doi.org/10.3758/s13423-017-1282-z

Lemaire, B., & Portrat, S. (2018). A Computational Model of Working Memory Integrating Time-Based Decay and Interference. *Frontiers in Psychology*, *9*, 416. https://doi.org/10.3389/fpsyg.2018.00416

Lewandowsky, S. (1999). Redintegration and Response Suppression in Serial Recall: A Dynamic Network Model. *International Journal of Psychology*, *34*(5–6), 434–446. https://doi.org/10.1080/002075999399792

Lewandowsky, S., & Oberauer, K. (2015). Rehearsal in serial recall: An unworkable solution to the nonexistent problem of decay. *Psychological Review*, *122*(4), 674–699. https://doi.org/10.1037/a0039684

Liu, Y. S., & Caplan, J. B. (2020). Temporal grouping and direction of serial recall. *Memory & Cognition*, *48*(7), 1295–1315. https://doi.org/10.3758/s13421-020-01049-x

McClelland, J. L., & Chappell, M. (1998). Familiarity Breeds Differentiation: A Subjective-Likelihood Approach to the Effects of Experience in Recognition Memory. *Psychological Review*, *105*(4), 724–760. https://doi.org/10.1037/0033-295X.105.4.734-760

Mizrak, E., & Oberauer, K. (2021). What Is Time Good for in Working Memory? *Psychological Science*, *32*(8), 1325–1337.

Neath, I., Saint-Aubin, J., & Surprenant, A. M. (2022). Semantic Relatedness Effects in Serial Recall But Not in Serial Reconstruction of Order. *Experimental Psychology*, 1618-3169/a000557. https://doi.org/10.1027/1618-3169/a000557

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411–421. https://doi.org/10.1037/0278-7393.28.3.411

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language*, *49*(4), 469–483. https://doi.org/10.1016/S0749-596X(03)00080-9

Oberauer, K. (2022). When does working memory get better with longer time? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(12), 1754–1774. https://doi.org/10.1037/xlm0001199

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic Bulletin & Review*, *18*(1), 10–45. https://doi.org/10.3758/s13423-010-0020-6

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*(5), 779–819. https://doi.org/10.3758/s13423-012-0272-4

Oberauer, K., & Lin, H.-Y. (2024). An interference model for visual and verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *in press*.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*(4), 761.

Penney, C. G. (1975). Modality effects in short-term verbal memory. *Psychological Bulletin*, *82*(1), 68–84. https://doi.org/10.1037/h0076166

Poirier, M., & Saint-Aubin, J. (1995). Memory for Related and Unrelated Words: Further Evidence on the Influence of Semantic Factors in Immediate Serial Recall. *The Quarterly Journal of Experimental Psychology Section A*, *48*(2), 384–404. https://doi.org/10.1080/14640749508401396

Popov, V., & Reder, L. M. (2020). Frequency effects on memory: A resource-limited theory. *Psychological Review*, *127*(1), 1–46. https://doi.org/10.1037/rev0000161

Portrat, S., Guida, A., Phénix, T., & Lemaire, B. (2016). Promoting the experimental dialogue between working memory and chunking: Behavioral data and simulation. *Memory & Cognition*, *44*(3), 420–434. https://doi.org/10.3758/s13421-015-0572-9

Portrat, S., & Lemaire, B. (2015). Is Attentional Refreshing in Working Memory Sequential? A Computational Modeling Approach. *Cognitive Computation*, *7*(3), 333–345. https://doi.org/10.1007/s12559-014-9294-8

Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: Insights from studies on attention and working memory: An overview of working memory consolidation. *Annals of the New York Academy of Sciences*, *1424*(1), 8–18. https://doi.org/10.1111/nyas.13633

Ricker, T. J., & Vergauwe, E. (2022). Boundary conditions for observing cognitive load effects in visual working memory. *Memory & Cognition*. https://doi.org/10.3758/s13421-022-01320-3

Ryan, J. (1969). Temporal Grouping, Rehearsal and Short-Term Memory. *Quarterly Journal of Experimental Psychology*, *21*(2), 148–155. https://doi.org/10.1080/14640746908400207

Saint-Aubin, J., & Poirier, M. (1999). Semantic Similarity and Immediate Serial Recall: Is There a Detrimental Effect on Order Information? *The Quarterly Journal of Experimental Psychology Section A*, *52*(2), 367–394. https://doi.org/10.1080/713755814

Saint-Aubin, J., Yearsley, J. M., Poirier, M., Cyr, V., & Guitard, D. (2021). A model of the production effect over the short-term: The cost of relative distinctiveness. *Journal of Memory and Language*, *118*, 104219. https://doi.org/10.1016/j.jml.2021.104219

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*(2), 145–166. https://doi.org/10.3758/BF03209391

Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review*, *15*(3), 535–542. https://doi.org/10.3758/PBR.15.3.535

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