‘NOT Gate’, Not ‘Gate’:

Long-term Memory Interplay with Working Memory via Reversal Coding

Shengyuan Wang (wangshy56@mail2.sysu.edu.cn)

Sun Yat-sen University, 510006

Yingchao Zhang

Sun Yat-sen University, 510006

Xiaowei Ding (dingxw3@mail.sysu.edu.cn)

Sun Yat-sen University, 510006

Abstract

Working memory (WM) is capacity-limited but can recruit long-term memory (LTM) to overcome such limitations. A critical question regarding this interplay is: How does LTM content flow into WM? Specifically, while recruiting LTM content that aligns with the current WM task is beneficial, recruiting inconsistent LTM content might impair performance. In this study, we investigated this question using EEG and the inverted encoding model to decode both WM and LTM content in real time. Contrary to the mainstream 'flexible gate' hypothesis, which suggests that a gate selectively allows beneficial LTM content to enter WM, we found that LTM content was consistently decoded, regardless of whether it was beneficial. Importantly, LTM content was represented in a reversed manner (inhibition) compared to WM content (activation). We supported this reversal coding of LTM functions to minimize interference with WM, benefiting WM performance. Our findings challenge the 'flexible gate' theory and suggest a 'NOT gate' mechanism regarding how LTM interplays with WM, where the coding of LTM content is systematically reversed after entering WM.

Keywords: working memory; long-term memory; EEG; inverted encoding model

Introduction

Working memory (WM) is capacity-limited (Luck & Vogel, 1997). To support complex cognitive functions, WM can recruit preexisting knowledge from long-term memory (LTM; Oberauer, 2009). However, indiscriminately incorporating all LTM content does not always benefit WM tasks. While recruiting LTM content that aligns with the current WM task is beneficial, recruiting inconsistent LTM content might impair performance. Therefore, a critical question about the interplay between LTM and WM is: How does information flow from LTM into WM?

One mainstream theory, the 'flexible gate' hypothesis (Mızrak & Oberauer, 2022; Oberauer et al., 2017), posits that a flexible gate selectively allows beneficial LTM content into WM while blocking harmful content.

In the present study, we employed EEG combined with the inverted encoding model (IEM; Sprague et al., 2014) to investigate how information flows from LTM into WM and critically tested the flexible gate hypothesis.

Method

Given that Experiments 1 and 2 employed identical paradigms with different stimulus sets, we focus on Experiment 1 (N = 35) here. It comprised three tasks:

IEM Training Task. Participants memorized six orientations. Using recorded EEG data, we trained an IEM to reconstruct orientation representations based on EEG channel responses. The inverted encoding model (IEM) was trained using data from all 64 EEG channels. Unlike traditional classification methods, IEM can decode multiple representations simultaneously. By reconstructing channel responses associated with distinct orientations, we computed the slope centered on a given orientation (Figure 1a &b). Specifically, we centered the profile on the target orientation and symmetrically folded the left side onto the right (Figure 1a & b), allowing for slope estimation relative to the target. A significant negative slope indicates activation, a significant positive slope indicates inhibition, and a non-significant slope suggests no representation. To validate the trained IEM, we applied it to the same IEM training task, confirming successful decoding through a significant negative slope (Figure 2a).

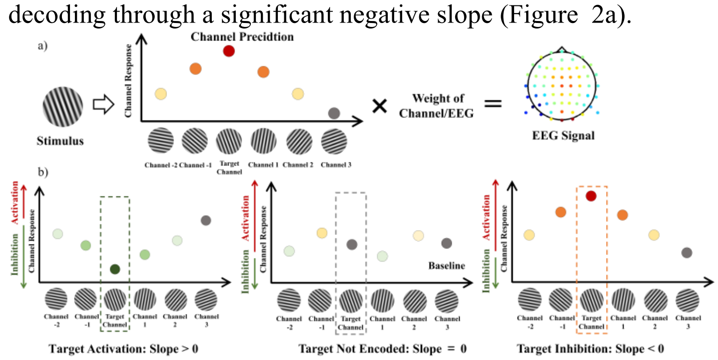

Figure 1: a) Training of IEM; b) Potential IEM outputs.

LTM Learning Task. Participants memorized three color-orientation pairs in LTM. This phase included two learning sessions followed by two testing sessions to ensure successful learning (Figure 2b).

WM Task. Participants sequentially learned two color-orientation pairs. A color cue was then presented, requiring them to recall the associated orientation (Figure 2c). To detect potential latent (or silent) WM/LTM representations, we applied a pinging impulse (Wolff et al., 2017), as such representations typically become accessible only after impulse stimulation.

The WM task included three conditions of color-orientation pairs (Figure 2c and d). In the consistent LTM condition, each color – orientation pair matched a previously learned association from the LTM task. Thus, retrieving the LTM pair could facilitate WM performance. In the inconsistent LTM condition, the color had been paired with an orientation in the LTM task but was now associated with a different orientation. As a result, retrieving the original LTM association would interfere with WM performance. In the no LTM condition, the color was novel and had not appeared in the LTM task, such that no prior association could be retrieved to affect the WM task.

We examined WM and LTM decoding during the encoding of the first color-orientation pair and the retrieval phase of the WM task. Under the flexible gate hypothesis, successful WM decoding but failed LTM decoding in the inconsistent LTM condition would be expected.

.

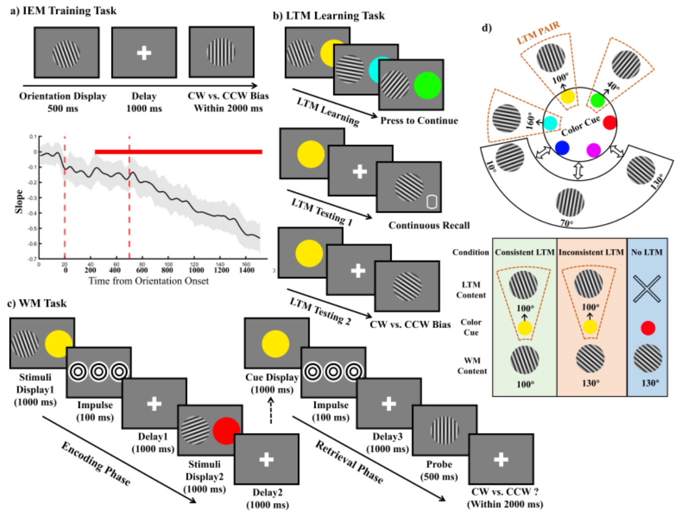

Figure 2: a-c) Procedure of IEM Training, LTM Learning, and WM Task; d) Three conditions of color-orientation pairs: consistent LTM condition, inconsistent LTM condition, and no LTM condition.

Results and Discussion

Contrary to the flexible gate hypothesis, LTM content was successfully decoded (cluster-based permutation) regardless of its consistency with WM content (both in the consistent and inconsistent LTM conditions).

Specifically, we computed IEM slopes separately for the orientation maintained in WM and the orientation previously associated with the same color cue in LTM, specifically within the inconsistent LTM condition. Additionally, we extracted IEM slopes for WM orientations in the no LTM condition, as well as for the shared WM–LTM orientation in the consistent LTM condition.

LTM orientations were consistently decodable in both the inconsistent and consistent LTM conditions, whereas WM orientations were decodable in the inconsistent and no LTM conditions (Figure 3a).

Critically, rather than being positively coded like WM content (activation: negative slope), LTM content was encoded in a reversed manner (inhibition: positive slope), as shown in Figure 3a.

To further validate reversal coding, we compared the averaged channel responses between positive (WM in No LTM condition) and reversal coding (LTM in Consistent LTM Condition). If reversal coding occurs, channel responses should be lower for reversal coding than for positive coding (Figure 3b). As predicted, reversal coding yielded significantly lower channel responses during retrieval (Figure 3c).

We propose that reversal coding serves a functional role: by inverting the sign of LTM representations, interference with WM is minimized, thereby enhancing WM performance. To test this, we divided participants into two groups based on WM performances. Supporting this view, participants with superior WM performance exhibited stronger reversal coding of LTM.

In a conceptual replication (Experiment 2), we replaced the separate color and Gabor (two objects) with a single colored Gabor while keeping all other settings identical to Experiment 1. The results replicated those of Experiment 1, further confirming their robustness.

Together, our findings challenge the 'flexible gate' theory and instead suggest that information flow from LTM to WM operates under a 'NOT gate' mechanism.

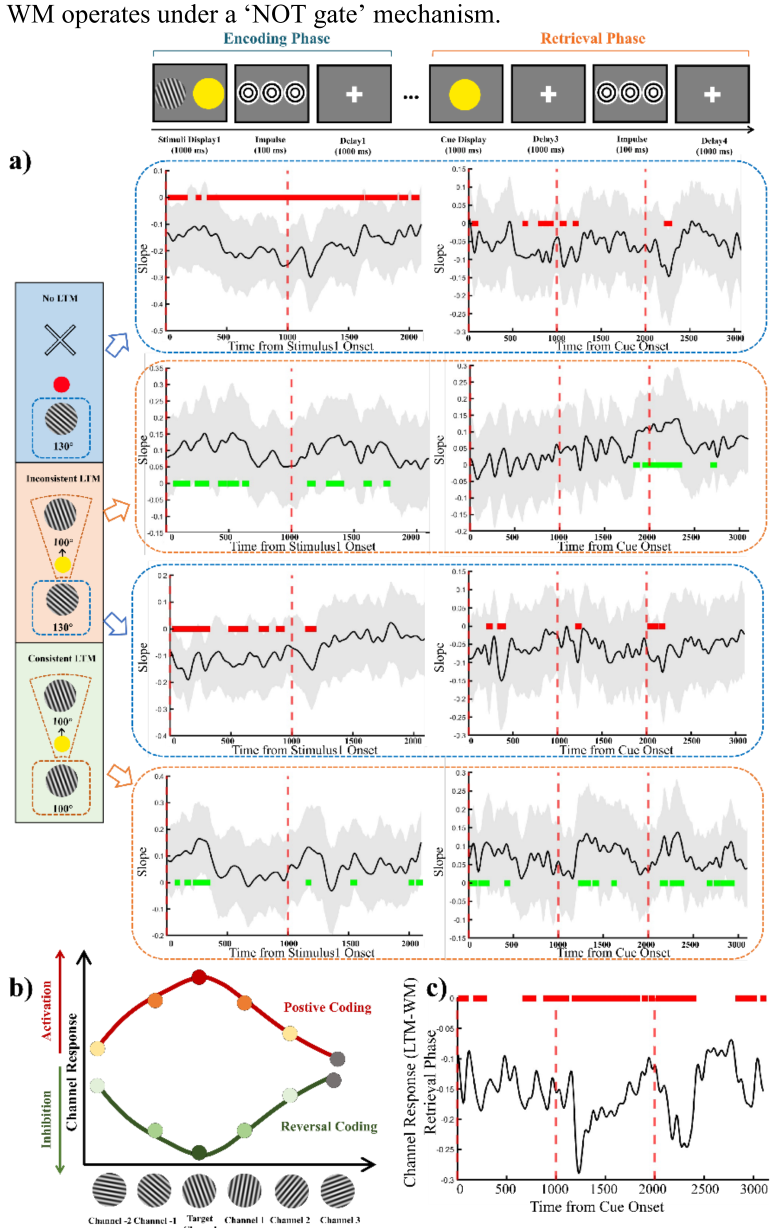

Figure 3: a) Decoding of WM and LTM contents across three conditions. Unlike WM content, which was positively coded (negative slope, with significant time duration marked by red lines), LTM content was encoded in a reversed manner (positive slope, with significant time duration marked by green lines); b) Diagram of positive coding and reversal coding; b) Differences in averaged channel responses between reversal (LTM) and positive (WM) coding. Red lines indicate the time duration during which reversal coding elicited significantly lower channel responses compared to positive coding.

Acknowledgments

This research was supported by grants from National Natural Science Foundation of China (32271103), Natural Science Foundation of Guangdong Province (2023A1515011762), Guangzhou Science and Technology Plan Project—Leading Elite Program (2024A04J3301), and the Fundamental Research Funds for the Central Universities, Major Project Cultivation and Emerging Interdisciplinary Cultivation Plan (24wkjc02) awarded to Xiaowei Ding.

References

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846.

Mızrak, E., & Oberauer, K. (2022). Working memory recruits long-term memory when it is beneficial: Evidence from the Hebb effect. *Journal of Experimental Psychology: General*, *151*(4), 763–780. https://doi.org/10.1037/xge0000934.

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), Psychology of learning and motivation (Vol. 51, pp. 45–100). *Elsevier*. https://doi.org/10.1016/S0079-7421(09)51002-X

Oberauer, K., Awh, E., & Sutterer, D. W. (2017). The role of long-term memory in a test of visual working memory: Proactive facilitation but no proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(1), 1-22. http://doi.org/10.1037/xlm0000302

Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of information in visual spatial working memory degrade with memory load. *Current Biology, 24*(18), 2174-2180. https://doi.org/10.1016/j.cub.2014.07.066.

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience, 20*(6), 864-871. https://doi.org/10.1038/nn.4546