WORKING MEMORY

Reactivation of latent working memories with transcranial magnetic stimulation

Nathan S. Rose,[1,2]* Joshua J. LaRocque,[1,3] Adam C. Riggall,[1,5] Olivia Gosseries,[1,4] Michael J. Starrett,[1] Emma E. Meyering,[1] Bradley R. Postle[1,5]*

The ability to hold information in working memory is fundamental for cognition. Contrary to the long-standing view that working memory depends on sustained, elevated activity, we present evidence suggesting that humans can hold information in working memory via "activity-silent" synaptic mechanisms. Using multivariate pattern analyses to decode brain activity patterns, we found that the active representation of an item in working memory drops to baseline when attention shifts away. A targeted pulse of transcranial magnetic stimulation produced a brief reemergence of the item in concurrently measured brain activity. This reactivation effect occurred and influenced memory performance only when the item was potentially relevant later in the trial, which suggests that the representation is dynamic and modifiable via cognitive control. The results support a synaptic theory of working memory.

The ability to mentally retain information in an accessible state, to manipulate it, and to use it to guide behavior is a critical building block for cognition. It has long been assumed that the neural basis for this working memory (WM) ability is elevated and persistent neuronal firing (1). This assumption has been called into question by recent proposals that information can be held in WM via synaptic mechanisms that do not require sustained, elevated brain activity (2–4).

Building on theoretical frameworks that information can be held in WM in one of several states of activation (5, 6), we recorded neural activity while participants performed a multi-step task in which two items were presented as memoranda for each trial. A cue indicated which item would be tested by the impending recognition memory probe, followed by the probe, then by a second cue, and then a second probe (Fig. 1). There was equal probability following the first cue, but not the second, that the uncued item might be needed for an ensuing memory judgment. This procedure moves the uncued item into a different state than the cued item, which, by definition, is in the focus of attention (7). Cognitive theories refer to the intermediate state of this unattended memory item (UMI) as "activated long-term memory" (LTM) (5, 6).

For experiment 1, multivariate pattern analysis (MVPA) showed evidence for an active representation of the UMI that dropped to baseline levels (Fig. 2) (7–9). This suggests that information in WM (but outside of focal attention) can be maintained in a latent state via mechanisms other than sustained, elevated activity. Although a similar drop-to-baseline pattern is observed when participants are instructed to drop information

[1]Department of Psychiatry, University of Wisconsin, Madison, WI 53706, USA. [2]Department of Psychology, University of Notre Dame, Notre Dame, IN 46556, USA. [3]Neuroscience Training Program, University of Wisconsin, Madison, WI 53706, USA. [4]Coma Science Group, University of Liège, 4000 Liège, Belgium. [5]Department of Psychology, University of Wisconsin, Madison, WI 53706, USA.
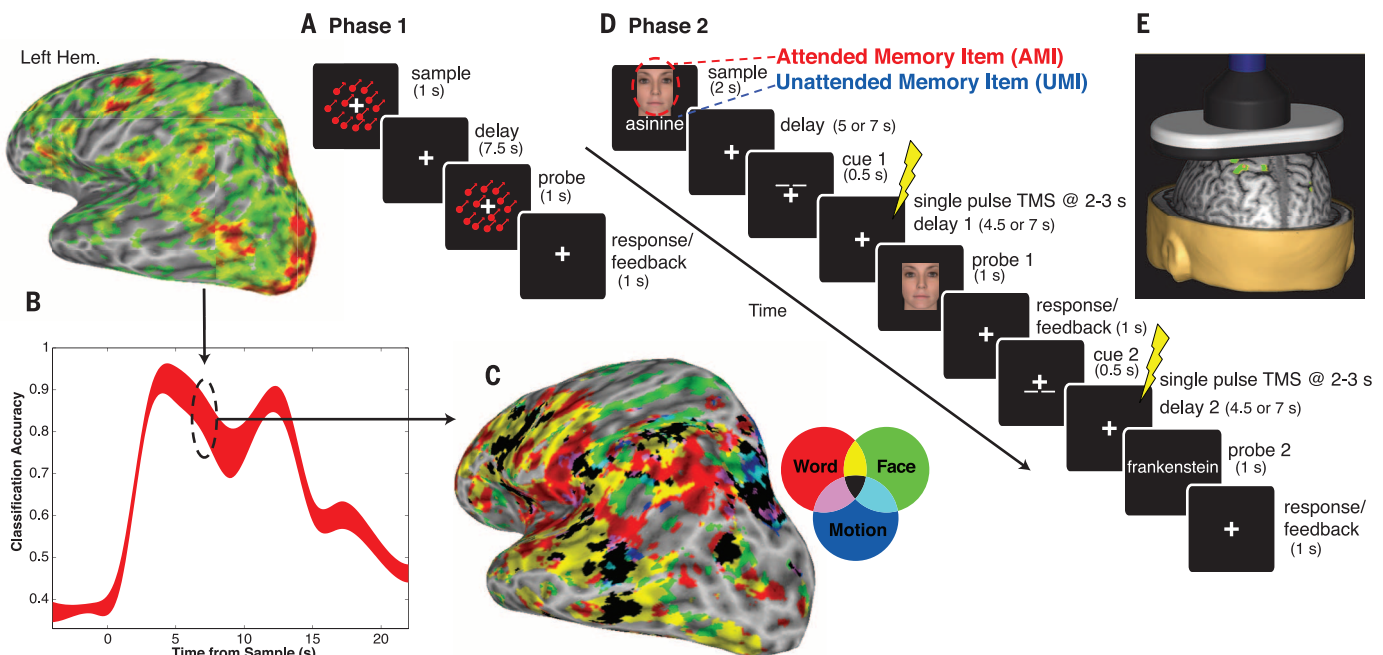*Corresponding author. Email: nrose1@nd.edu (N.S.R.); postle@wisc.edu (B.R.P.)

Fig. 1. General procedure. (**A**) In phase 1, functional magnetic resonance imaging (fMRI) data were acquired while participants performed a one-item delayed-recognition task for words, faces, or directions of motion; these data were used for multivariate pattern analysis (MVPA). (**B**) Classifiers trained on the delay period were used for subsequent analyses. For experiment 1, these classifiers were used to decode fMRI activity from phase 2 (Fig. 2). (**C** and **D**) For experiments 2 and 3, they were used in a whole-brain searchlight conjunction analysis to generate participant-specific maps of category-sensitive areas (C); nonoverlapping areas were used for transcranial magnetic stimulation (TMS) targeting in phase 2 (D). (**E**) In phase 2, single pulses of TMS were delivered during the postcue delay periods.

from WM (*10*, *11*), here the UMI remained in WM because, when so instructed by the second cue, participants accurately reactivated it and used it to evaluate the final probe (Fig. 2B).

In three additional experiments, we tested the hypothesis that if a UMI is encoded in a distributed pattern of synaptic weights and held in a state that is more accessible than trial-irrelevant information, the readout from a nonspecific burst of activity filtered through this network might reveal this latent representation (*2*) (fig. S1). This would be consistent with the idea that networks in the posterior cortex can be dynamically configured as matched filters to encode behaviorally relevant information (*3*, *4*, *12*, *13*).

For experiments 2 and 3, participants performed the phase 2 WM task (Fig. 1) while we recorded electroencephalography (EEG) and applied single-pulse transcranial magnetic stimulation (TMS) 2 to 3 s after the cue. For experiment 2, we targeted brain regions identified from the phase 1 MRI task as preferentially supporting MVPA decoding for one category, but not the other two. MVPA of the spectrally transformed EEG data from only the phase 2 task detected reliable evidence for an active representation of both memory items across the initial portion of the trial, until the onset of the first cue, at which point decoding accuracy remained elevated for the attended memory item (AMI) but dropped to the baseline for the UMI (*14*).

After a single pulse of TMS, there was a brief recovery of MVPA decoding of the UMI—a "reactivation effect"—before it returned to baseline and remained there while the cued item was tested [$P = 0.01$; Bayes factor (BF) = 3.64 against the null] (Fig. 3A). TMS affected neither broadband decoding of the AMI nor recognition memory judgments (fig. S4). When we analyzed bandpass-filtered data, the TMS reactivation effect was isolated to signal from the beta band (fig. S5) and was associated with a transient period of above-chance decoding performance for both the UMI and the AMI. The TMS reactivation effect was specific for information that was in WM on that trial, because above-chance MVPA performance, as assessed with the AUC (area under the curve) analysis, necessarily means that TMS did not activate a representation of the category that was irrelevant on that trial.

In experiment 2, we administered blocks of trials with TMS targeting one of the category-selective regions, but we varied, on a trial-by-trial basis, which category was the AMI and which was the UMI. Each block included trials for which the UMI belonged to the targeted region's preferred category, and trials for which it did not. A TMS reactivation effect was observed (Fig. 3B) whether or not TMS targeted the UMI's category-preferred region, although the effect was larger and more prolonged when it did (BF = 4.02 for targeted sites, 1.72 for nontargeted sites). This finding suggests that WM is supported by heightened connectivity between cortical networks that represent all trial-relevant information (AMI and UMI) relative to trial-irrelevant information (*15*, *16*).
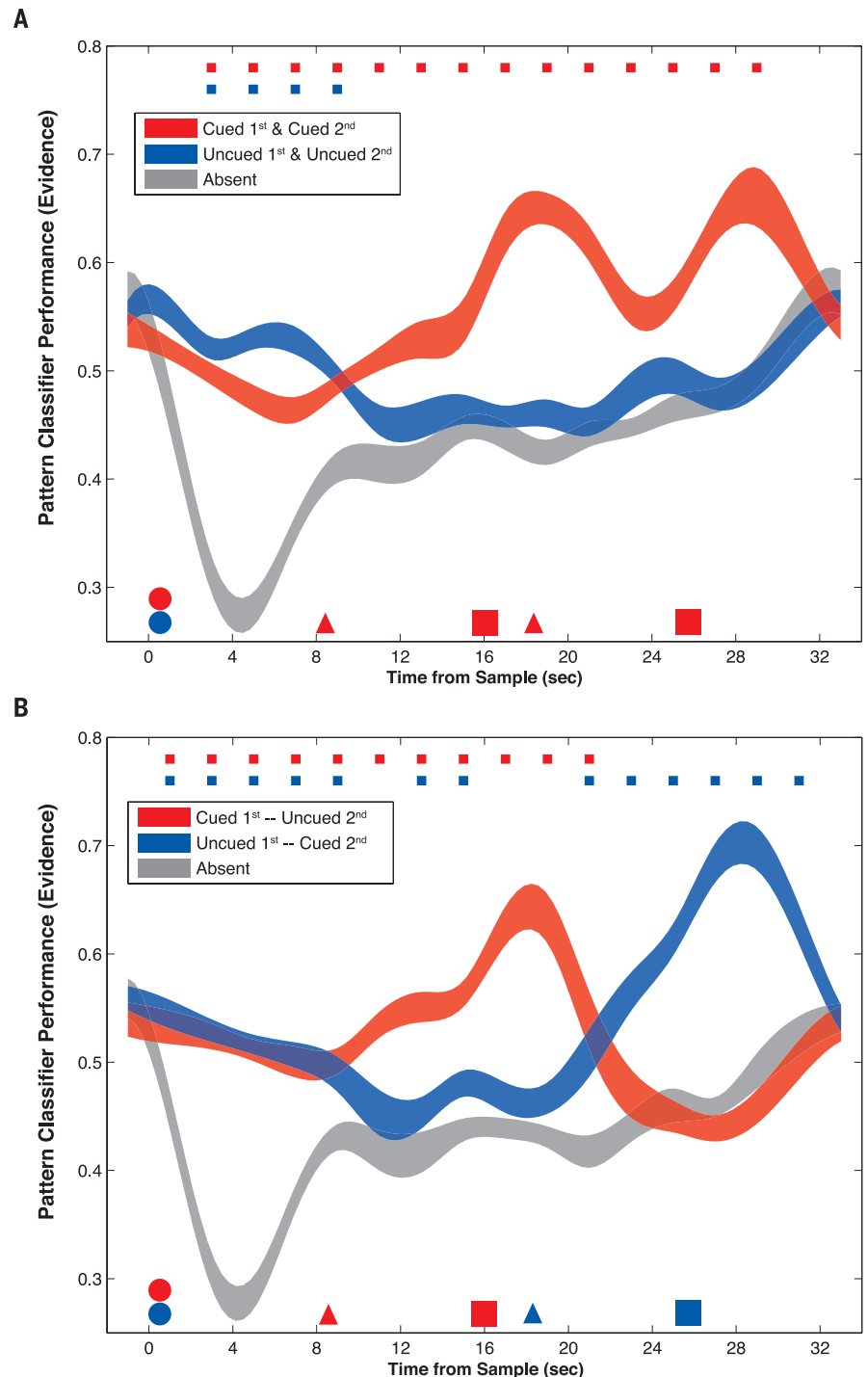

Fig. 2. Experiment 1 fMRI decoding (train phase 1, test phase 2): Classifier evidence as a function of an item's status, collapsed across stimulus category. After stimulus presentation (red and blue circles), delay-period classifier evidence for both items was elevated relative to the empirical baseline of evidence for the category that was not presented on that trial ("absent," gray). Upon presentation of the first cue (red triangle), evidence for the cued category (red) remained elevated, but for the uncued category (blue) dropped to baseline. (**A**) After the first probe (red square), on half the trials the second cue designated that the same item would be tested by the second probe, and evidence for the two categories remained the same relative to baseline. (**B**) When the second cue designated the previously uncued item, evidence for the two categories reversed for the remainder of the trial. Color-coded small squares at the top of each plot indicate $P < 0.01$; line width reflects SEM.
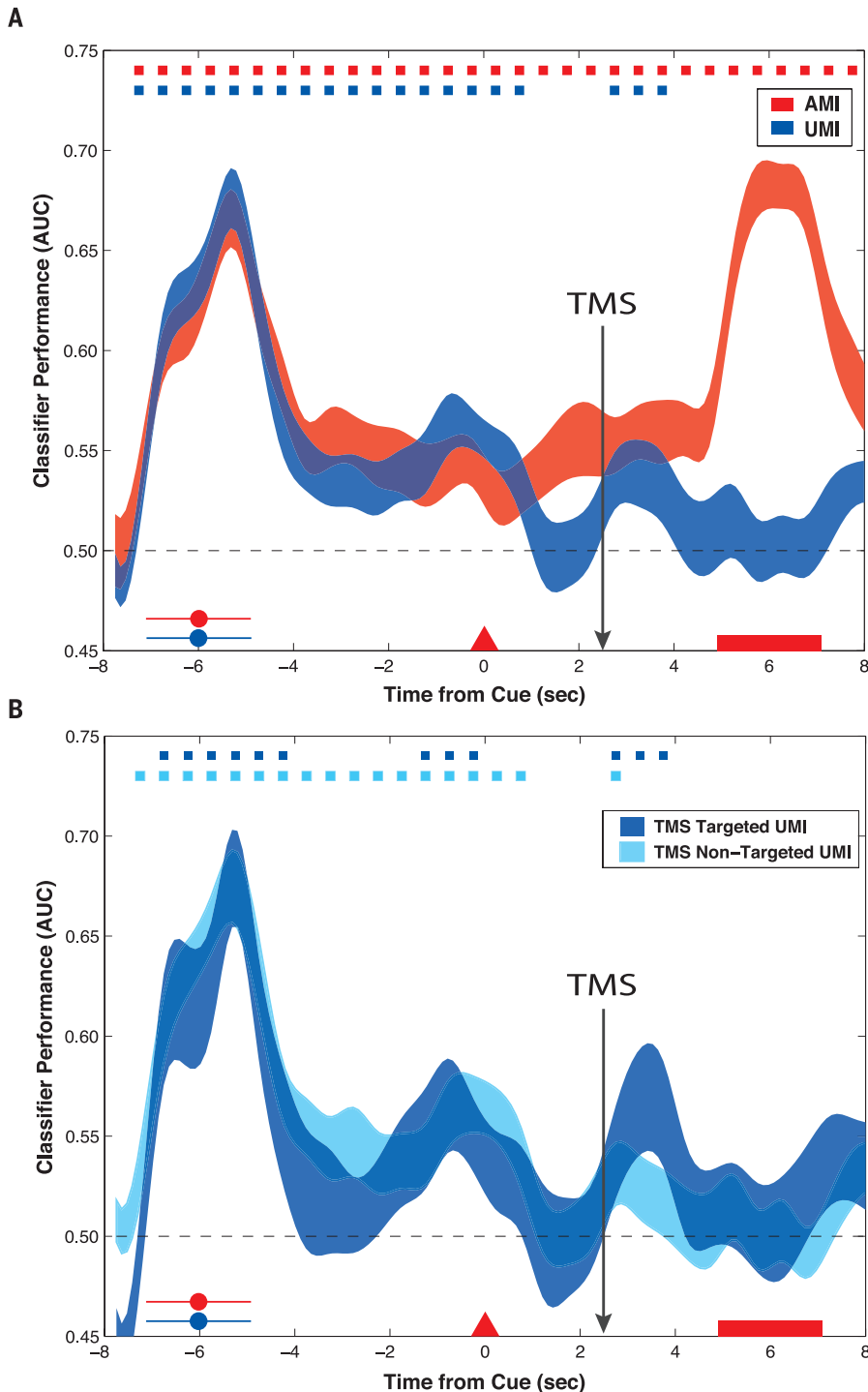
Fig. 3. Experiment 2 EEG decoding (train and test on phase 2 data): Classifier accuracy (area under curve, AUC) as a function of an item's status at the time of the first cue, collapsed across stimulus category. AUC reflects classifier sensitivity to discriminating between evidence for the AMI or UMI relative to the absent category. (**A**) Classification time series of the AMI and UMI upon stimulus presentation (red and blue circles), the first cue (red triangle), TMS, and first probe (red rectangle), averaged over N = 18 sessions, 2952 trials (decoding ends where the AMI and UMI switched on 50% of the trials). (**B**) Decoding UMIs as a function of whether TMS targeted that item's phase 1–defined region or a different category's region. Color-coded small squares at the top of each plot indicate P < 0.05; line width reflects SEM.

Retrocues that inform subjects that they can drop an item from memory result in a rapid loss of multivariate evidence for the no longer relevant item (11, 17). Nonetheless, proactive interference from stimuli presented on previous trials indicates that the brain retains a residual trace of such recent, but no longer relevant, information (18). An important test of state-based models of WM is whether there is a functional distinction between UMIs (putatively held in a state of activated LTM) and dropped information (no longer in WM). In experiment 3, with a different group of participants, we also administered TMS after the second cue, after which the uncued item would no longer be relevant on the trial, and at which point it should have the same status as an irrelevant item. If the TMS reactivation effect is a consequence of an item being maintained in a privileged state, it should only be observed when that item is still potentially relevant for the trial. We also jittered the onset of TMS between 2 and 3 s after the cues (14) and standardized TMS by targeting the same region on every trial for all participants—an MVPA-defined region in the right precuneus known to be critical for the top-down control of visual attention (19) (Fig. 4A).

For the first half of the trial, the results from experiment 3 replicated those from experiment 2 (Fig. 4B), with a robust TMS reactivation effect for the UMI (BF = 9.8 against the null). For the delay period following the second cue, however, there was no evidence for significant decoding of the uncued item following the TMS pulse (BF = 3.4 in favor of the null). These results suggest that UMIs are maintained in a different state than are items that have been dropped from WM, and that the mechanisms that maintain latent representations in WM are dynamic and modifiable via cognitive control (20).

Because our design entails decoding at the category level, it does not rule out the possibility that the TMS reactivation effect reflects a general reinstatement of category context (21), rather than the temporary activation of the UMI itself. The idea that the representation of the UMI itself drives this effect would be strengthened by demonstrating that TMS can influence recognition memory decisions on this task. If the TMS reactivation effect reflects a temporary reinstatement of the UMI back into the focus of attention, participants should have more difficulty rejecting the UMI as a lure when probing their memory of the AMI.

In experiment 4, we presented recognition memory probes that matched the AMI on 50% of trials; of the 50% of nonmatch probes, 30% were drawn from the same category as the AMI, and a critical 20% matched the UMI (14). Participants were instructed to reject memory probes that did not match the AMI. Critically, only for the first probe was there an increased proportion of false alarms to the UMI for TMS relative to no-TMS trials (Fig. 3C, P = 0.01, BF = 3.48) (14).

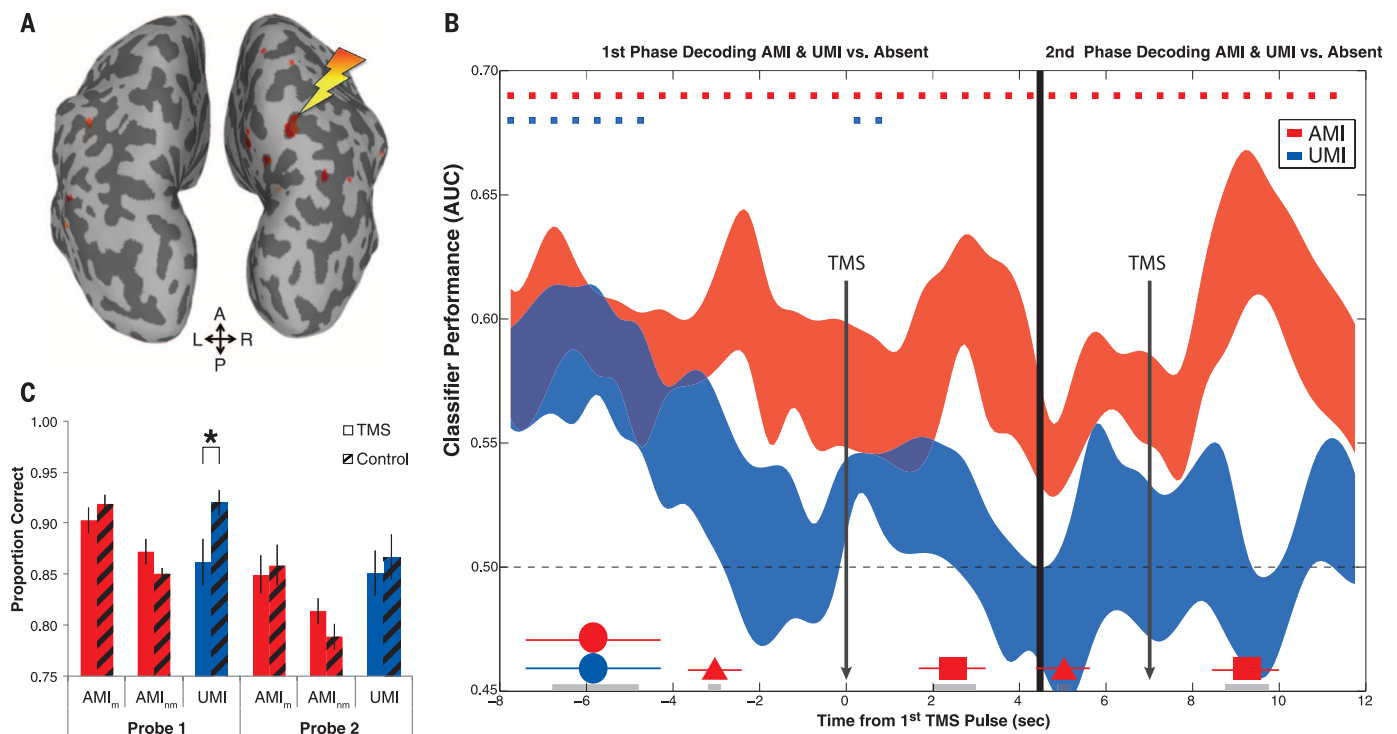Our results have important implications for the understanding of WM at many levels. They

Fig. 4. Results from experiments 3 and 4. (**A**) The MVPA-defined TMS target for experiments 3 and 4 (right precuneus). A, anterior; L, left; R, right; P, posterior. (**B**) Classification time series from experiment 3 showing TMS reactivation of the UMI after the first cue, when the UMI was still relevant (left), but not after the second cue, when the UMI was no longer relevant on the trial (right) averaged over 1152 trials. Color-coded small squares at the top of each plot indicate $P < 0.05$; line width reflects SEM. (**C**) Experiment 4 recognition memory for AMI match probes (AMI$_m$), AMI nonmatch probes (AMI$_{nm}$), and UMI (nonmatch) probes. *$P < = 0.01$; error bars denote SEM.

provide neural evidence for at least two levels of WM that are distinct from the default state of LTM representations (5, 6). They are inconsistent with models positing just one level of WM storage (22, 23). They also suggest that instead of "activated LTM," a more apt label for the second level of WM would be "prioritized LTM." Information can be held in WM in latent "activity-silent" traces (11, 20). What might be the physiological bases of such representations? Computational models of WM have proposed that short-term synaptic plasticity could be the basis for the transient formation of weight-based networks that can represent information over short time periods (2, 24).

Our results provide empirical evidence for the existence of a short-term plasticity mechanism that is likely to be fundamental to a wide range of cognitive functions involving attentional selection (25) and may provide the building blocks for long-term potentiation mechanisms that support LTM (26). Therefore, our findings introduce a potential avenue for reactivating and strengthening representations that underlie many classes of high-level cognition.

REFERENCES AND NOTES

1. P. S. Goldman-Rakic, *Neuron* **14**, 477–485 (1995).
2. G. Mongillo, O. Barak, M. Tsodyks, *Science* **319**, 1543–1546 (2008).
3. Y. Sugase-Miyamoto, Z. Liu, M. C. Wiener, L. M. Optican, B. J. Richmond, *PLOS Comput. Biol.* **4**, e1000073 (2008).
4. M. J. Wolff, J. Ding, N. E. Myers, M. G. Stokes, *Front. Syst. Neurosci.* **9**, 123 (2015).
5. N. Cowan, in *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, A. Miyake, P. Shah, Eds. (Cambridge Univ. Press, 1999), pp. 62–101.
6. K. Oberauer, *Psychol. Learn. Motiv.* **51**, 45–100 (2009).
7. J. J. LaRocque, J. A. Lewis-Peacock, A. T. Drysdale, K. Oberauer, B. R. Postle, *J. Cogn. Neurosci.* **25**, 127–142 (2013).
8. J. A. Lewis-Peacock, A. T. Drysdale, K. Oberauer, B. R. Postle, *J. Cogn. Neurosci.* **24**, 61–79 (2012).
9. J. J. LaRocque, A. C. Riggall, S. M. Emrich, B. R. Postle, *Cereb. Cortex* 10.1093/cercor/bhw283 (2016).
10. E. F. Ester, D. E. Anderson, J. T. Serences, E. Awh, *J. Cogn. Neurosci.* **25**, 754–761 (2013).
11. A. C. Riggall, B. R. Postle, *J. Neurosci.* **32**, 12990–12998 (2012).
12. B. Y. Hayden, J. L. Gallant, *Front Neurosci* **7**, 18 (2013).
13. N. E. Myers *et al.*, *eLife* **4**, e09000 (2015).
14. See supplementary materials on *Science* Online.
15. B. Kundu, J. Y. Chang, B. R. Postle, B. D. Van Veen, *Neuroimage* **114**, 320–327 (2015).
16. T. G. Lee, M. D'Esposito, *J. Neurosci.* **32**, 15458–15466 (2012).
17. T. B. Christophel, M. N. Hebart, J. D. Haynes, *J. Neurosci.* **32**, 12983–12989 (2012).
18. S. Monsell, *Cognit. Psychol.* **10**, 465–501 (1978).
19. D. M. Beck, N. Muggleton, V. Walsh, N. Lavie, *Cereb. Cortex* **16**, 712–717 (2006).
20. M. G. Stokes, *Trends Cogn. Sci.* **19**, 394–405 (2015).
21. S. M. Polyn, V. S. Natu, J. D. Cohen, K. A. Norman, *Science* **310**, 1963–1966 (2005).
22. B. McElree, *Psychol. Learn. Motiv.* **46**, 155–200 (2006).
23. A. Baddeley, *Annu. Rev. Psychol.* **63**, 1–29 (2012).
24. V. Itskov, D. Hansel, M. Tsodyks, *Front. Comput. Neurosci.* **5**, 40 (2011).
25. R. Desimone, J. Duncan, *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
26. Y. Dudai, *Annu. Rev. Psychol.* **55**, 51–86 (2004).

Science

Reactivation of latent working memories with transcranial magnetic stimulation

Nathan S. Rose, Joshua J. LaRocque, Adam C. Riggall, Olivia Gosseries, Michael J. Starrett, Emma E. Meyering and Bradley R. Postle

How to reactivate forgotten memories

Sophisticated techniques can decode stimulus representations for items held in a person's working memory. However, when subjects shift their attention toward something else, the neural representation of the now unattended item drops to baseline, as though the item has been forgotten. Rose *et al.* used single-pulse transcranial magnetic stimulation (TMS) to briefly reactivate the representation of an unattended item. A short pulse of TMS enhanced recognition of "forgotten" stimuli, bringing an unattended item back into focal attention.
Science, this issue p. 1136

ARTICLE TOOLS	http://science.sciencemag.org/content/354/6316/1136
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2016/12/01/354.6316.1136.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/354/6316/1089.full http://stm.sciencemag.org/content/scitransmed/5/198/198ra105.full http://stm.sciencemag.org/content/scitransmed/5/200/200ra115.full http://stm.sciencemag.org/content/scitransmed/7/278/278ra33.full http://stm.sciencemag.org/content/scitransmed/8/330/330re1.full
REFERENCES	This article cites 36 articles, 7 of which you can access for free http://science.sciencemag.org/content/354/6316/1136#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service