

---

## S1 Supplemental

### S1.1 High vs. Low Confidence EEG Clusters

Here trials were selected based on the confidence ratings given for each retrieval attempt with the intention of capturing differences in EEG time frequency for Recollection vs. Familiarity based memory retrieval. The same analysis approach used in the closed-loop vs. open-loop results was adopted for this confidence contrast (i.e. cluster based search in Stimulus Locked and Response Locked within theta, alpha and beta frequency bands). All correctly retrieved trials with confidence rating less than 3 were assigned to the Low Confidence condition, and all correctly retrieved trials with confidence greater than 2 were assigned to the High Confidence condition (trial counts: low confidence correct  $\mu = 25 \pm 4$  95% confidence interval across subjects, high confidence correct  $\mu = 127 \pm 33$ ). It should be noted that this difference in trial counts per condition should not impact the type 1 error rate of the cluster statistics as significance is calculated through a bootstrapped random permutation of condition labels across subjects.

Figure S1 shows the only 3 clusters found using the  $p < 0.05$  criteria; all clusters were found in the Response Locked epochs. The theta cluster showed a greater increase in power relative to baseline for the High Confidence condition compared to Low Confidence, had a temporal extent from -3000 ms to -680 ms Response Locked, and its  $t$ -value mass was centered in left posterior electrodes. In contrast, the alpha cluster showed a greater decrease in power for the Low Confidence condition compared to High Confidence, had a temporal extent from -3000 ms to -480 ms Response Locked, and its  $t$ -value mass was in central posterior electrodes. Similarly the beta cluster showed a greater decrease in power for the Low Confidence condition compared to High Confidence, had a temporal extent from -3000 ms to -1600 ms Response Locked, and its  $t$ -value mass was centered in right anterior electrodes.

### S1.2 Theta-Phase Model

Input patterns were constructed, and memory retrieval performance measured, based on the slot topology in the EC layers (as highlighted in Figure 2). This slot structure is intended to capture the modality segregation within EC, and within each slot we assume there is a *vocabulary* of different patterns, which reflect the representational repertoire within those modalities. These vocabulary patterns were similarly used to estimate error within the networks' output by comparing, within a given slot, the output pattern of activation with all other vocabulary patterns. If, for the given input pattern, the slots' output at the  $EC_{out}$  layer is closest to the vocabulary pattern it was trained on, it is considered correct, and otherwise considered incorrect. This closest-pattern calculation is done across each of the slots for every input pattern, and if any slot shows an incorrect response the network output for that input pattern is counted as incorrect. This measurement is referred to as *Name Error* in the results section, and is thought to better represent the potential for clean up of hippocampal output as compared to more standard measures such as Sum Squared Error (SSE). It also has the advantage of not requiring any further threshold or other parameterization. It should be noted that this measure of error, compared to a SSE, deemphasizes single unit based errors in output in favor of an emphasis on distributed patterns of error across groups of units.

## S2 Attentional Modulation: Deep LEABRA

The ‘Deep LEABRA’ framework currently being developed in the Emergent software is the initial explorations into the modeling of interactions between thalamic relay nuclei, and superficial vs. deep cortical layers (O'Reilly, Wyatte, & Rohrlich, 2014). This framework proposes that the neocortex is composed of two separable networks: superficial and deep/thalamic. The superficial-layer network consists of neocortical layers 4, 2, and 3, across different brain areas. The deep/thalamic network starts in each cortical area with the layer 5b intrinsic bursting (IB) neurons, which receive inputs from local

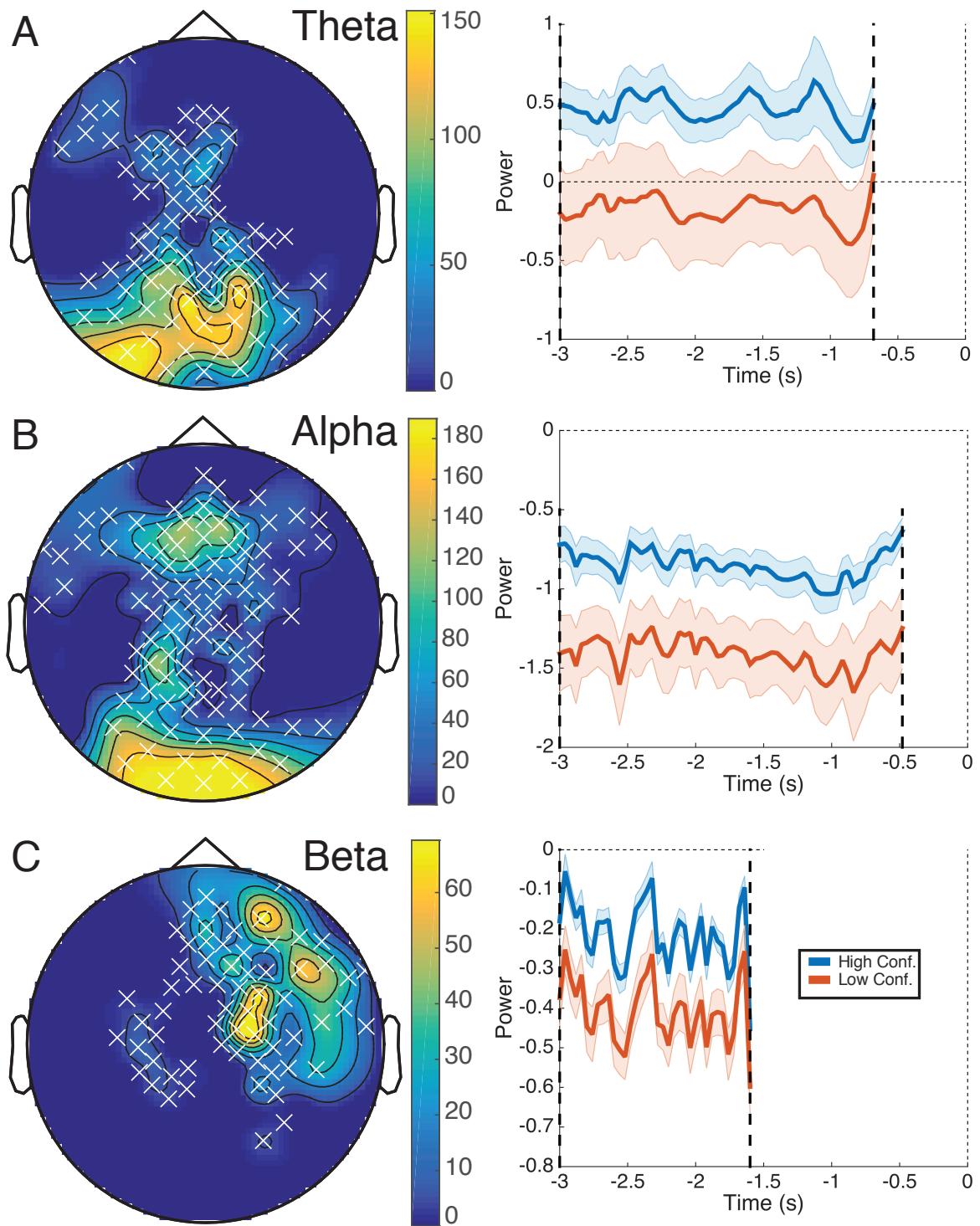


FIGURE S1: (A) Clustered found for response locked power difference between High vs. Low Confidence correct retrieval trials in the theta band. On the left is the summed  $t$ -values over time for significant electrodes, on the right is the average power across significant electrodes for High Confidence shown in blue and Low Confidence shown in red. Large dashed lines show the boundaries of the significant temporal window. (B) Alpha cluster found for same contrast as in A, here Low confidence trials show a greater decrease from baselined compared to high confidence trials. (C) Beta cluster found, with similar properties to the alpha cluster in B.

---

superficial neurons and top-down projections from other areas, and provide strong driving feedforward input to higher-order thalamic nuclei. These 5IB neurons also project to deep layer 6 cortical-thalamic (6CT) neurons, which interconnects with the thalamus (which in turn projects back up to layer 4 of the superficial network and layer 6 in the deep network). This connectivity is illustrated in Figure S2. The current work takes advantage of the biological referents and computational implementation of the Deep LEABRA framework however, here it is used as a simple form of attentional modulation (and not a temporal learning mechanism as described in O'Reilly et al. (2014)). The minimal description of the components necessary for this implementation are described below.

The modulation of standard LEABRA unit activations (referred to as superficial units here) based on the Deep LEABRA framework and the thalamic gating model described in N. A. Ketz, Jensen, and O'Reilly (2015) is carried out in two basic steps. First, a Thalamic Relay Cell (TRC) unit receives input based on the threshold filtered output of some upstream units which have received its input from processing a given stimulus (in this case the hippocampus is providing this top-down modulation based on what is being retrieved and sent to entorhinal cortex). These sending units, for a given TRC, are assumed to be the cortical thalamic projections from neocortical layers 5IB, 6CT (Sherman & Guillery, 2006), which reflect a thresholded filtering of superficial layer activity from both top-down and bottom-up sources. TRC activation dynamics are calculated in a similar manner to all LEABRA units, described in detail in (N. Ketz, Morkonda, & O'Reilly, 2013), with the exception that its inputs are subject to an activation threshold of 0.1, and the weighted connections into and out of the TRC are not learned but instead remain fixed at 0.8.

Second, the TRC activation is used to calculate the attentional modulation that it sends to down stream superficial units. The attentional modulation is implemented through a separate set of activations for receiving units that can be considered an attentional mask that gets applied. This mask is the TRC activation normalized to a range between some pre-determined minimum and a maximum of 1. The effective modulation of the receiving unit is simply the multiplication of this attentional mask, and the receiving units' net input. Thus, where activations are strong in the mask, the corresponding superficial layer activations will remain strong, but where they are weaker, the superficial layer activations will be reduced. This modulation process is hypothesized to be carried out by layer 6CT neurons (Bortone, Olsen, & Scanziani, 2014; Olsen, Bortone, Adesnik, & Scanziani, 2012).

This multiplicative modulation is based off previous mathematical models of attention that capture many of the effects relevant to the perception literature (Reynolds & Heeger, 2009; Montijn, Klink, & Van Wezel, 2012). Figure S2 shows how this framework captures the essential computations of the Reynolds and Heeger (2009) model in different parts of the superficial and deep layer circuits. Computationally, the TRC activations encode the attentional modulations of the superficial layer state. The attentional modulation signals cause the iterative constraint satisfaction process in the superficial network to focus on task-relevant information while down-regulating responses to irrelevant information, consistent with the abstract Reynolds and Heeger (2009) model. Its important to note that this process can work in both a top-down (shown in S2) and bottom-up fashion. Biologically, the layer 6CT neurons are known to exhibit a multiplicative influence over firing of superficial-layer neurons, in a manner consistent with the Reynolds and Heeger (2009) model (Bortone et al., 2014; Olsen et al., 2012).

In this work the minimum TRC activation was set to 0.25 through a loose process of trying to minimize the number of training epochs and matching the pattern of human behavioral performance. This process consisted of iteratively training networks each with progressively higher minimum TRC modulation values, and finding the network that produced the best paired associate training accuracy with the smallest number of training epochs. As stated this was loosely fit and a more precise value could be arrived at with more systematic methods.

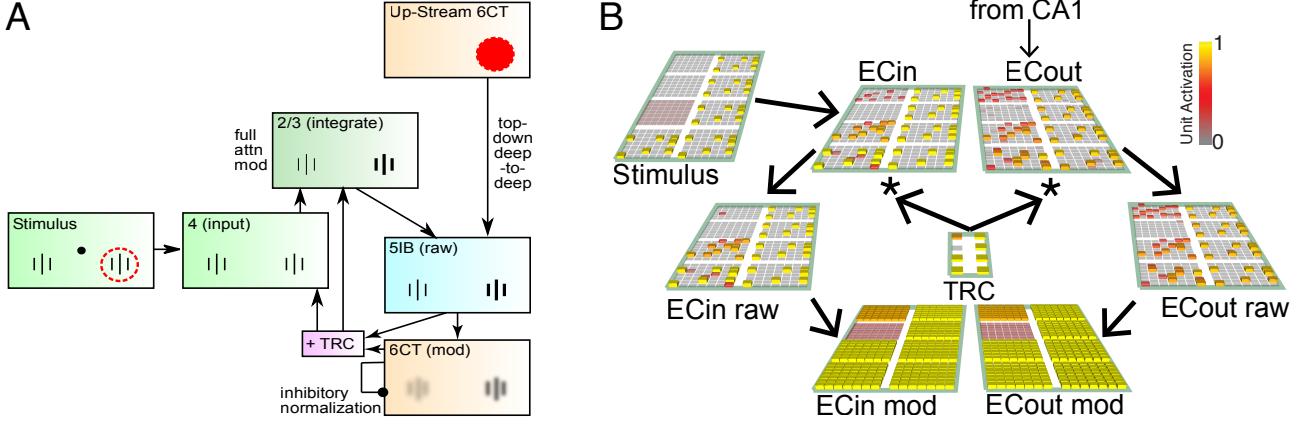


FIGURE S2: How attentional modulation is computed in response to a top-down attentional focus. (A) Layer 4 receives bottom-up sensory input, which then drives superficial layers (2/3), which initially do not reflect the attentional modulation (not shown). The deep layer 5 Intrinsic Bursting (5IB) neurons integrate deep-to-deep top-down attentional inputs plus the local stimulus features from 2/3, to produce a *raw* deep output, prior to the contextual normalization process. Layer 6 Cortical-Thalamic (6CT) neurons then integrate this direct activation from 5IB, to produce, for the first time in the circuit, a properly renormalized multiplicative gain-field activation pattern, with surround inhibition both within the 6CT layer and further downstream in the TRC circuit providing the renormalization process. These 6CT activations then modulate (multiply) the superficial-layer activations to produce *both* an increase in the attended location, and a decrease for the unattended location, as shown. (B) Implementation of computational principle within the Theta-Phase network. In the model layer 4 is subsumed into layer 2/3 neurons, as represented by ECin and ECout. Top down influence comes from the hippocampus via area CA1 connections to ECout. Layer 5IB *raw* signals are calculated from ECin and ECout, and in turn used to calculate the *mod* signal. This layer 6CT project to the TRC units which ultimately are used to modulate the EC superficial layers with the current attentional mask.

### S2.1 Dependent Events Example Stimuli

Below is an example set of events generated for the Dependent Events paradigm. These stimuli were originally adapted from Horner, Bisby, Bush, Lin, and Burgess (2015), where some original elements were removed due to British vs. American ubiquity, and replaced with more common American elements. The first 18 events are open loop, and the last 18 are closed loop with the appropriate balancing of object and animal stimuli.

Event Number	Location	People	Objects	Animals
1	train station	George Clooney	football	cat
2	supermarket	Britney Spears	wallet	dog
3	cinema	Mel Gibson	hammer	hamster
4	restaurant	Beyonce	lightbulb	lion
5	bar	Justin Bieber	battery	peacock
6	office	Hilary Clinton	skateboard	monkey
7	parking lot	Vladimir Putin	necklace	slug
8	gym	Angelina Jolie	violin	chicken
9	bowling alley	George Bush	notebook	sheep
10	elevator	Bill Murray	handbag	cow
11	tree house	Barack Obama	stroller	ferret
12	living room	Jim Carrey	trophy	goldfish
13	hospital	Ellen DeGeneres	shovel	lobster
14	church	Robert De Niro	bike	moose
15	swimming pool	Bono	calculator	rat
16	cruise ship	Christopher Walken	carrot	panda
17	coffee shop	Tom Hanks	camera	octopus
18	zoo	Stephen Hawking	book	duck

19	nightclub	Sean Connery	trumpet	
20	park	Johnny Depp	basket	
21	school yard	Taylor Swift	umbrella	
22	police station	Paris Hilton	mug	
23	beach	Lady Gaga	television	
24	stadium	Emma Watson	baseball bat	
25	ski lift	Bill Gates	magnet	
26	kitchen	Donald Trump	mirror	
27	bedroom	Madonna	whisk	
28	basement	Clint Eastwood		fox
29	patio	Paul McCartney		frog
30	airport	Harrison Ford		goat
31	highway	Oprah Winfrey		scorpion
32	river	John Travolta		badger
33	corn field	Scarlett Johansson		parrot
34	bank	Pamela Anderson		dolphin
35	hair salon	Meryl Streep		wolf
36	casino	Angela Merkel		giraffe

## References

- Bortone, D. S., Olsen, S. R., & Scanziani, M. (2014, April). Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex. *Neuron*, 82.
- Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, 6, 7462. doi:10.1038/ncomms8462
- Ketz, N. A., Jensen, O., & O'Reilly, R. C. (2015, January). Thalamic pathways underlying prefrontal cortex-medial temporal lobe oscillatory interactions. *Trends in neurosciences*, 38, 3–12.
- Ketz, N., Morkonda, S. G., & O'Reilly, R. C. (2013, June). Theta coordinated error-driven learning in the hippocampus. *PLoS Computational Biology*, 9, e1003067.
- Montijn, J. S., Klink, P. C., & Van Wezel, R. J. A. (2012). Divisive normalization and neuronal oscillations in a single hierarchical framework of selective visual attention. *Frontiers in Neural Circuits*, 6, 22. doi:10.3389/fncir.2012.00022
- Olsen, S., Bortone, D., Adesnik, H., & Scanziani, M. (2012, February). Gain control by layer six in cortical circuits of vision. *Nature*, 483(7387), 47–52.
- O'Reilly, R. C., Wyatte, D., & Rohrlich, J. (2014, July). Learning Through Time in the Thalamocortical Loops. *arXiv:1407.3432 [q-bio]*. arXiv: 1407.3432 [q-bio]
- Reynolds, J. H. & Heeger, D. J. (2009, January). The normalization model of attention. *Neuron*, 61(2), 168–185.
- Sherman, S. & Guillery, R. (2006). *Exploring the Thalamus and Its Role in Cortical Function*. Cambridge, MA: MIT Press.