

# How Hippocampus and Cortex Contribute to Recognition Memory: Revisiting the Complementary Learning Systems Model

Kenneth A. Norman\*

**ABSTRACT:** We describe how the Complementary Learning Systems neural network model of recognition memory (Norman and O'Reilly (2003) *Psychol Rev* 104:611–646) can shed light on current debates regarding hippocampal and cortical contributions to recognition memory. We review simulation results illustrating three critical differences in how (according to the model) hippocampus and cortex contribute to recognition memory, all of which derive from the hippocampus' use of *pattern separated* representations. Pattern separation makes the hippocampus especially well-suited for discriminating between studied items and related lures; it makes the hippocampus especially poorly suited for computing global match; and it imbues the hippocampal ROC curve with a  $Y$ -intercept  $> 0$ . We also describe a key boundary condition on these differences: When the average level of similarity between items in an experiment is very high, hippocampal pattern separation can fail, at which point the hippocampal model will start to behave like the cortical model. We describe the implications of these simulation results for extant debates over how to describe hippocampal versus cortical contributions and how to measure these contributions. © 2010 Wiley-Liss, Inc.

**KEY WORDS:** computational methods; neural networks; recognition memory

## INTRODUCTION

Over the past decade, there has been a heated debate over how to characterize the contributions of the hippocampus vs. medial temporal lobe cortex to recognition memory. Some researchers have argued that hippocampus and cortex implement recollection and familiarity, respectively [e.g., Yonelinas (2002) and Aggleton and Brown (1999)], whereas others have argued that this distinction does not adequately capture the differences in these structures' contributions [e.g., Squire et al. (2007)]. One reason that this debate has lasted so long is because of disagreement over how to measure hippocampal and cortical contributions. Yonelinas and colleagues have developed several techniques for estimating recollection and familiarity based on behavioral data (Yonelinas, 1994; Yonelinas et al., 1998); these procedures rely on a core set of assumptions about the properties of recollection and familiarity. If these assumptions adequately describe the (respective) contributions of hippocampus versus cortex to recognition memory, then Yonelinas' measurement procedures, which were originally developed apart from neural considerations, can

be used to estimate the contributions of these structures. However, the validity of the assumptions is still very much in question [e.g., Wixted (2007) and Wixted and Squire (2008)].

In this work, we describe how a computational model that my colleagues and I have developed [the Complementary Learning Systems (CLS) neural network model; Norman and O'Reilly (2003)] can help to address these controversies. Importantly, this model does not build in any specific claims about how hippocampus versus cortex might contribute to recognition memory. Rather, our approach has been to start out with some widely accepted ideas about key anatomical and physiological properties of hippocampus and cortex, and then use our computational models to explore the implications of these ideas for how these structures might contribute to recognition judgments. The most important of these ideas is that hippocampus implements *pattern separation* (i.e., it is biased to assign distinct representations to input patterns, regardless of their similarity) but cortex does not. The model allows us to explore the implications of this difference in pattern separation for recognition performance.

In the Complementary Learning Systems Framework section of this work, we review the basic Complementary Learning Systems account of the division of labor between hippocampus and cortex (McClelland et al., 1995). In the Applying CLS to Recognition Memory Data section, we describe how we have applied our models of hippocampus and cortex to recognition memory data. In the Key Results from our CLS Recognition Memory Simulations section, we review simulation results illustrating three critical differences in how (according to the model) hippocampus and cortex contribute to recognition memory; all these differences relate to the difference in pattern separation described earlier. We also describe a key boundary condition (i.e., limit) on these differences: When the average level of similarity between items in an experiment is very high [e.g., all of the stimuli in the experiment are computer-generated male faces; Yotsumoto et al. (2007)], hippocampal pattern separation can fail, at which point the hippocampal model will start to behave like the cortical model. Finally, in the Implications for the Current Debate section, we describe the implications of these simulation results for extant debates about the neural basis of recognition memory.

Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey

Grant sponsor: NIH; Grant numbers: NIH R01 MH069456.

\*Correspondence to: Kenneth Norman, Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540. E-mail: knorman@princeton.edu

Accepted for publication 17 June 2010

DOI 10.1002/hipo.20855

Published online in Wiley Online Library (wileyonlinelibrary.com).

## THE COMPLEMENTARY LEARNING SYSTEMS FRAMEWORK

The CLS framework developed by McClelland et al. (1995) incorporates several widely held ideas about the division of labor between hippocampus and neocortex that have been developed over many years by many different researchers [e.g., Scoville and Milner (1957), Marr (1971), Grossberg (1976), O’Keefe and Nadel (1978), Teyler and Discenna (1986), McNaughton and Morris (1987), Sherry and Schacter (1987), Rolls (1989), Sutherland and Rudy (1989), Squire (1992), Eichenbaum et al. (1994), Treves and Rolls (1994), Burgess and O’Keefe (1996), Wu et al. (1996), Moll and Miikkulainen (1997), Hasselmo and Wyble (1997), Aggleton and Brown (1999), and Becker (2005)]. According to the CLS framework, neocortex forms the substrate of our internal model of the structure of the environment. In contrast, hippocampus is specialized for rapidly and automatically memorizing patterns of cortical activity, and so they can be recalled later based on partial cues (*pattern completion*). The model posits that neocortex learns incrementally; each training trial results in relatively small adaptive changes in synaptic weights. These small changes allow cortex to gradually adjust its internal model of the environment in response to new information. The other key property of neocortex (according to the model) is that it assigns similar (overlapping) representations to similar stimuli. Use of overlapping representations allows cortex to represent the shared structure of events and therefore makes it possible for cortex to generalize to novel stimuli based on their similarity to previously experienced stimuli. In contrast, the model posits that hippocampus assigns distinct, *pattern separated* representations to stimuli, regardless of their similarity. This property allows hippocampus to rapidly memorize arbitrary patterns of cortical activity without suffering unacceptably high (catastrophic) levels of interference.

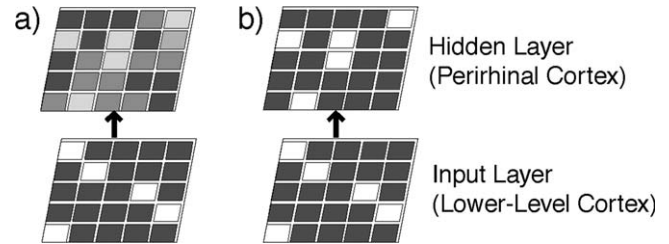
## APPLYING CLS TO RECOGNITION MEMORY DATA

Norman and O’Reilly (2003) built cortical and hippocampal neural networks that instantiate the CLS principles outlined above. Here, we describe how we applied these neural networks to recognition memory data.

### Cortical Contributions to Recognition Memory

The cortical model cannot support pattern completion of novel associations after a single study exposure (due to its relatively low learning rate), but it can discriminate between studied and nonstudied items based on the sharpness of their representations in perirhinal cortex.

Figure 1 illustrates the structure of the CLS cortical model (note that the actual model had more simulated neurons). The



**FIGURE 1.** Illustration of the sharpening of hidden (perirhinal) layer activity patterns in a miniature version of the CLS cortical model. (a) Shows the network before sharpening; perirhinal activity (more active = lighter color) is relatively undifferentiated. (b) Shows the network after Hebbian learning and inhibitory competition produce sharpening; a subset of the units are strongly active, while the remainder are inhibited. Figure adapted from Norman and O’Reilly (2003).

model consists of an input layer (corresponding to lower regions of the cortical hierarchy), which projects in a feed-forward fashion to a hidden layer (corresponding to perirhinal cortex). As mentioned earlier, the main function of cortex is to extract statistical regularities in the environment; the two-layer CLS cortical network (where perirhinal hidden units compete to encode regularities that are present in the input layer) is meant to capture this idea in the simplest possible fashion. Because the cortical model uses a small learning rate, it is not capable of pattern completion following limited exposure to a stimulus. However, it is possible to extract a scalar signal from the cortical model that discriminates between previously presented and novel stimuli. In the cortical model, as items are presented repeatedly, their representations in the upper (perirhinal) layer become *sharper*. Novel stimuli weakly activate a large number of perirhinal units, whereas previously presented stimuli strongly activate a relatively small number of units. Sharpening occurs in the model because Hebbian learning specifically tunes some perirhinal units to represent the stimulus. When a stimulus is first presented, some perirhinal units (by chance) will respond more strongly to the stimulus than other units. These “winning” units get tuned by Hebbian learning to respond even more strongly to the item than next time it is presented; this increased response triggers an increase in feedback inhibition to units in the layer, resulting in decreased activation of the “losing” units. This latter property (whereby some initially responsive units drop out of the stimulus representation as it is repeated) is broadly consistent with the neurophysiological finding that some perirhinal neurons show decreased responding as a function of how often a stimulus has been presented [e.g., Xiang and Brown (1998), Li et al. (1993), and Viskontas et al. (2006)]. Figure 1 illustrates this sharpening dynamic.

In the Norman and O’Reilly (2003) paper, the cortical memory signal was operationalized by reading out the activation of the  $k$  winning units in the perirhinal layer (where  $k$  is a model parameter that defines the maximum number of units that are allowed to be strongly active at once), although other methods of operationalizing familiarity are possible [e.g., we

can measure how quickly activation spreads to the perirhinal layer of the network from the input layer; Norman and O'Reilly (2003)]. We should also note that sharpening is not a novel property of the CLS model. Rather, it is a general property of competitive-learning networks with graded unit-activation values in which there is some kind of contrast enhancement within a layer [see, e.g., Grossberg (1986), Section 23, and Grossberg and Stone (1986), Section 16].

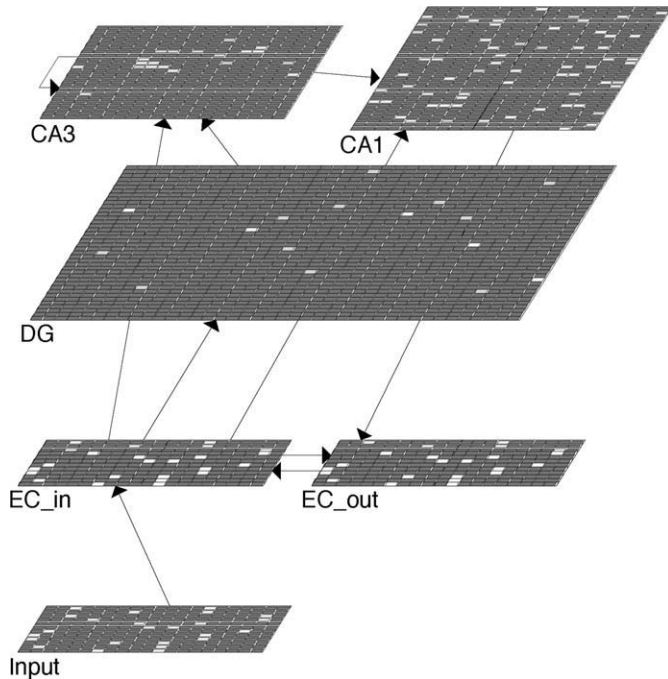
Although the Norman and O'Reilly's (2003) cortical model provides a straightforward implementation of sharpening, the model falls short in several ways. For example, Bogacz and Brown (2003) demonstrated that the simple Hebbian learning rule used in the model yields grossly inadequate storage capacity (i.e., the number of novel vs. familiar patterns that can be successfully discriminated does not scale well as a function of network size). The model also fails to account for cortex's ability to complete well-learned patterns. Extant evidence from amnesic patients suggests that, while the hippocampus may be needed to complete novel patterns that have only been studied once, patients with relatively spared medial temporal lobe cortex can learn to recall novel associations (e.g., new semantic facts) after extensive exposure to these novel associations [e.g., Gardiner et al. (2008)]. Because the Norman and O'Reilly (2003) cortical model only has feedforward connections (i.e., it lacks feedback connections from the upper layer to the lower layer, and it lacks recurrent connections within layers), the model lacks the ability to fill in missing pieces of well-learned input patterns. To remedy these deficits, we have explored a variant of the model that includes feedback and recurrent connections and uses a new learning algorithm [the *Oscillating Learning Algorithm*; Norman et al. (2005, 2006)]. The new version of the model has much better storage capacity than the original version, and it also has the ability to fill in missing pieces of stored patterns after extensive exposure to those patterns. Importantly, all the cortical model properties that are described in Key Results from our CLS Recognition Memory Simulations section (below) also apply to this updated version of the model.

Another shortcoming of the Norman and O'Reilly (2003) cortical model is that it does not directly address questions of functional specialization within cortex: In our description of the model (above), we asserted that that judgments of the familiarity of object stimuli are based on the sharpness of representations in perirhinal cortex, but we did not explain *why* perirhinal cortex (out of all cortical regions) plays this role. With regard to this question, the CLS framework (as originally formulated by McClelland et al.) is broadly compatible with the representational-hierarchical view of cortical organization expressed by Cowell et al. (this issue). According to this view, cortex is arranged into a hierarchy of layers, where each layer learns to represent conjunctions of features on the layer below it. Perirhinal cortex represents highly complex conjunctions of object features by virtue of its position at the top of the ventral visual stream hierarchy. As described by Cowell et al. [this issue; see also Bussey et al. (2002), Bussey and Saksida (2002), Barense et al. (2005), and Cowell et al. (2006, 2010)], this ability to represent complex feature conjunctions makes perirhi-

nal cortex especially useful for high-level object perception (i.e., discriminating between objects that contain different combinations of complex features); it also makes perirhinal cortex especially useful for object recognition memory, compared to regions lower in the hierarchy. The simple visual features (e.g., straight lines) that are represented in lower cortical regions are likely to be present in both studied items and lures, rendering them useless for discriminating between studied items and lures on a recognition test; by contrast, the more complex visual features that are represented in perirhinal cortex are more diagnostic (i.e., they are less likely to be shared across studied items and lures), thereby making them more useful for discriminating between studied items and lures (Cowell et al., 2006, this issue).

The main difference between the CLS framework and the representational-hierarchical framework relates to the degree of conjunctivity attributed to representations in perirhinal cortex. Cowell et al. (2006, this issue) posit that each layer in the cortical hierarchy learns conjunctions of features in the preceding layer, regardless of how useful (or useless) these conjunctive representations are for predicting meaningful outcomes. This process culminates in perirhinal neurons that represent conjunctions at the "whole object" level, such that different objects (i.e., objects that differ in at least one feature) are assigned non-overlapping representations in perirhinal cortex. The CLS view of conjunctive learning in cortex is somewhat different; this view is most fully described in a paper by O'Reilly and Rudy (2001) that incorporated error-driven learning into the cortical model (in addition to Hebbian learning). The key idea from O'Reilly and Rudy (2001) is that, while cortex has the ability to develop conjunctive representations, it does not learn these representations automatically—rather, it acquires conjunctive representations in a task-driven fashion [i.e., whenever the conjunctive representation is needed to reduce prediction error; see also O'Reilly and Norman (2002)]. Because conjunctive representations are learned on an "as needed" basis, representations in higher-level cortical regions in the O'Reilly and Rudy (2001) model tend to be relatively less conjunctive and more overlapping than the representations posited by Cowell et al. (2006, this issue). We should emphasize that this difference is largely a matter of degree (both frameworks posit the existence of conjunctive representations in perirhinal cortex). Nonetheless, we think that the difference is consequential: The CLS framework predicts that representations in perirhinal cortex should overlap to the extent that objects share features, whereas the framework posited by Cowell et al. does not. This "graded representational similarity" property is very important for the global match predictions described in the coming section.

One last issue relates to the difference between perirhinal and parahippocampal contributions to recognition memory. Although this difference has not been explicitly modeled in our work thus far, we should note that the CLS framework is broadly compatible with the BIC framework described by Diana et al. (2007) and Ranganath (this issue). We believe that the same basic computational operation (hierarchical feature extraction) is taking place in ventral stream structures (perirhinal cortex) and dorsal stream structures (parahippocampal



**FIGURE 2.** Diagram of the CLS hippocampal network. The hippocampal network links input patterns in entorhinal cortex (EC) to relatively nonoverlapping (pattern-separated) sets of units in region CA3. The dentate gyrus (DG) serves to facilitate pattern separation in region CA3. Recurrent connections in CA3 bind together all the units involved in representing a particular EC pattern; the CA3 representation is linked back to EC via region CA1. Learning in the CA3 recurrent connections, and in projections linking EC to CA3 and CA3 to CA1, makes it possible to recall entire stored EC patterns based on partial cues. Figure reprinted from Norman and O’Reilly (2003).

cortex); the only difference is in the types of information that are represented. Ventral-stream structures represent object information, making this stream useful for judging the familiarity of individual objects; dorsal-stream structures represent other types of information (e.g., relating to spatial configuration) that can be used to contextualize object representations.

### Hippocampal Contributions to Recognition Memory

The hippocampal model, unlike the cortical model, can support recall (pattern completion) of stored patterns after a single study exposure. This allows the model to make recognition decisions based on the content of retrieved information, not just the amount of retrieved information.

Figure 2 shows the hippocampal model. To apply the hippocampal model to recognition, Norman and O’Reilly (2003) compared the test cue (presented on the EC\_in layer, which corresponds to superficial layers of entorhinal cortex) to the pattern of retrieved information (activated over the EC\_out layer, which corresponds to deep layers of entorhinal cortex). When recalled information matches the test cue, this constitutes evidence that the item was studied; conversely, mismatch between recalled information and the test cue constitutes evi-

dence that the test cue was not studied (e.g., study “rats,” test “rat”; if the hippocampal model recalls that rats-plural was studied, not rat-singular, this can serve as grounds for rejection of rat). For every item, we computed a hippocampal “memory evidence” score by taking the number of recalled features that matched the test cue and subtracting out the number of recalled features that mismatched the test cue. Note that none of the simulation results described below depend on our use of this *match–mismatch* measure. For example, the properties described below are still present when we use the average activity of the hippocampal network (in region CA1 of the model) as our measure of memory evidence.

Importantly, although the CLS model provides a detailed account of the properties of the hippocampal and cortical memory signals, it does not directly address the question of how subjects use these signals to make recognition memory decisions. In the Norman and O’Reilly (2003) model, recognition decision-making was treated as a “black box” that was external to the network model itself (i.e., the decision process was not itself simulated by a neural network). The simulations from that paper used a simple decision rule whereby test items were called “old” if the hippocampal signal exceeded a certain value; otherwise, the decision was made based on the cortical signal (Jacoby et al., 1997). This decision rule is almost certainly oversimplified [e.g., we agree with Wixted (2007) that subjects probably combine the signals in some way before making their decision, instead of treating the two signals separately]. Developing a richer account of recognition memory decision rules and the neural mechanisms that implement these rules is an important direction for future research [see Norman et al. (2008) for further discussion of these issues].

## KEY RESULTS FROM OUR CLS RECOGNITION MEMORY SIMULATIONS

In this section, we review three critical differences in how (according to our model) hippocampus and cortex contribute to recognition memory. We then describe an important boundary condition on these differences.

### Critical Difference 1: Differences in How the Cortical and Hippocampal Signals Are Distributed

#### *Distributional properties of the cortical memory signal*

The memory signal generated by the cortical network is well described by a Gaussian signal detection model. The ratio of studied variance to lure variance depends on which sources of variability are present in the paradigm being simulated.

Figure 3 shows the distributions of memory scores associated with studied items and lures from a representative cortical-model simulation. Both these distributions have an approximately Gaussian shape (note that this is an emergent property of the model rather than something that was built into the model in the first place). In Figure 3, the distributions are shown as hav-

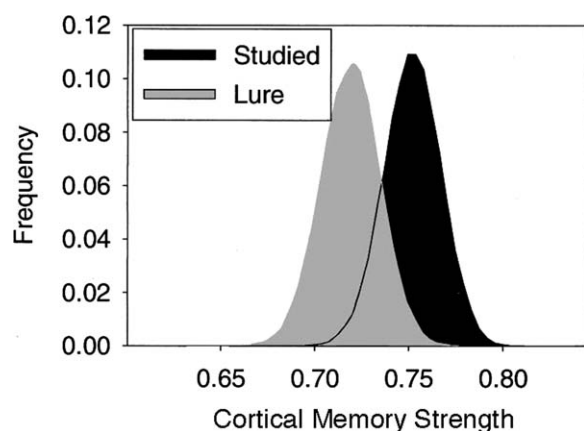


FIGURE 3. Histogram of the memory signal generated by the CLS cortical network, given 20% average overlap between input patterns. Figure adapted from Norman and O'Reilly (2003).

ing equal variance, but this is not always the case. As discussed in Norman and O'Reilly (2003), the ratio of the variances will depend on the particular sources of variability that are present in a particular simulation. Some kinds of variability, for example, encoding variability (variability in how well items are stored during the study phase), affect studied items more than lures and should therefore lead to greater variability in the studied-item memory strength distribution than in the lure memory strength distribution (Hilford et al., 2002; Wixted, 2007). Other forms of variability, for example, pre-experimental variability (variability in how much the features of a test item overlap with features of items studied before the experiment) should affect studied items and lures equally. If the ratio of pre-experimental variability to encoding variability is high (i.e., encoding variability is making a relatively small contribution to performance), then studied-item variance and lure variance will be approximately equal. Conversely, if encoding variability is more prominent than pre-experimental variability in a particular simulation, then variance in cortical memory strength (across items) will be larger for studied items than for lures.

### *Distributional properties of the hippocampal memory signal*

The memory signal generated by the hippocampal network is both continuous and non-Gaussian. The maximal level of confirmatory evidence generated by studied items typically exceeds the maximal level of evidence generated by lures. This property results in a positive  $Y$  intercept on ROC curves.

Figure 4 shows [for a representative simulation taken from Norman and O'Reilly (2003)] the distributions of hippocampal memory scores associated with studied items and lures. This particular simulation had a moderate level of feature overlap between input patterns (20% on average). As is clear from Figure 4, the hippocampal memory signal distributions do not adhere to a simple Gaussian model. The bulk of the lure distribution is located at the zero point, although some lures trigger above-zero memory signal. The studied distribution is bimodal, and crucially, it extends further to the right than the lure distri-

bution, so there are some (high) memory scores that are sometimes triggered by studied items but never by lures.

The low level of lure memory strength shown in Figure 4 is a consequence of hippocampal pattern separation. In the hippocampal model, recall (pattern completion) occurs when the test item activates the hippocampal neurons that were involved in representing an item at study. Because of pattern separation, the hippocampal representations of lures and studied items typically do not overlap strongly; this low level of overlap prevents lures from accessing stored hippocampal memory traces and triggering recall. Because the hippocampal memory strength distribution for studied items extends past the hippocampal memory strength distribution for lures, the model predicts that the hippocampal memory process will generate an ROC curve with a positive  $Y$ -intercept (i.e., it should be possible to place a decision boundary so there are some hits but no false alarms).

These properties of the CLS hippocampal model are broadly consistent with the idea (promoted by Yonelinas and colleagues) that the hippocampus generates a *thresholded* memory signal that is sometimes triggered by studied items but never by lures (Yonelinas, 1994; Yonelinas et al., 1998). If we treat the upper limit of the lure distribution in Figure 4 as a "threshold," it is true that studied items sometimes exceed the threshold but lures never do. Above-threshold memory strength values (in Fig. 4, memory strength values  $>0.2$ ) are perfectly diagnostic as to whether the test item as studied (because these high values are only ever generated by studied items), whereas below-threshold memory strength values carry very little information about whether the test item was studied. To be clear, the idea that the CLS hippocampal memory signal is a threshold process [in the technical, signal-detection sense; Macmillan and Creelman (2005)] does not imply that the hippocampus generates a discrete, all-or-none signal, and it does not imply that qualitatively different processes are engaged on trials where memory exceeds (vs. does not exceed) the threshold. In the CLS model, studied items and lures both trigger continuously varying levels of evidence; it just happens to be the case that (because of pattern separation) the maximum level of evidence

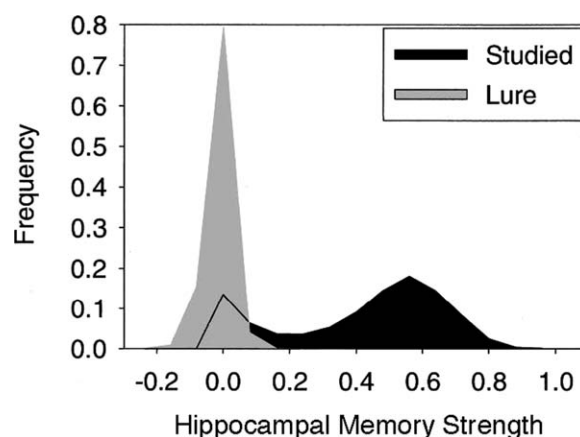


FIGURE 4. Histogram of the memory signal generated by the CLS hippocampal network, given 20% average overlap between input patterns. Figure adapted from Norman and O'Reilly (2003).

triggered by studied items exceeds the maximum level of evidence triggered by lures.

It is also useful to clarify how the CLS model relates to the memory model recently proposed by Shimamura and Wickens (2009). Like our model, the Shimamura and Wickens's (2009) model posits that the hippocampus can generate a non-Gaussian, positively skewed memory signal distribution, whereas regions lower in the MTL processing hierarchy generate less-skewed (more Gaussian) memory signal distributions. A crucial difference between the Shimamura and Wickens (2009) model and the CLS model is that the CLS model is a *mechanistic* model (i.e., it specifies a procedure for computing memory strength scores for individual simulated items), whereas the Shimamura model is a *measurement* model (i.e., it specifies a procedure for estimating model parameters from behavioral summary statistics, but it does not specify a procedure for computing memory strength scores for individual items). In the CLS model, the non-Gaussian shape of the hippocampal distribution is an emergent property of assumptions that we have made about how the hippocampal memory signal is computed; in the Shimamura model, the non-Gaussian shape of the hippocampal distribution is captured by their use of ex-Gaussian distributions, which contain a parameter for positive skew.

### **Critical Difference 2: The Hippocampus Performs Better Than Cortex at Discriminating Between Studied Items and Related Lures (Except If the Test Gives a Forced Choice Between Studied Items and Corresponding Related Lures)**

The CLS model predicts that cortex and hippocampus can both support good recognition performance when lures are not closely related to studied items. However, when lures are closely related to specific studied items [e.g., study “rats” and test with “rat”; Hintzman et al. (1992)], hippocampally based recognition performance should be higher than cortically based recognition performance, because of the hippocampus' ability to assign distinct representations to similar stimuli. As mentioned earlier, this pattern-separation property implies that most lures will not trigger a strong hippocampal signal. Furthermore, if a lure is similar enough to a studied item to trigger pattern completion of that item, the lure can be rejected based on mismatch between the test probe and the recalled information.

The CLS model also predicts that effects of target-lure similarity should interact with test format. Most recognition tests use a yes–no (YN) format where test items are presented one at a time, and subjects are asked to label them as old or new. The model predicts that cortex should perform very poorly on YN tests with related lures (because the distributions of cortical memory scores associated with studied items and related lures overlap strongly). However, the model predicts that cortex should perform much better when given a forced choice between studied items and corresponding related lures (e.g., “rat” and “rats” are presented simultaneously, and subjects have to choose which item was studied). In this situation, the model predicts

that the mean difference in cortical memory strength between the studied item and the related lure will be small, but the studied item should reliably trigger a slightly larger cortical memory signal than the corresponding related lure (thereby allowing for correct responding). This prediction comes about because there is extensive covariance in the cortical memory scores triggered by studied items and corresponding related lures; this high level of covariance makes the variance of the difference in these cortical memory scores very small [see Hintzman (1988) for further discussion of this point about covariance].

Taken together, these predictions imply that patients with hippocampal damage should perform very poorly on YN tests with related lures. However, the same patients should show relatively spared performance on tests with unrelated lures, or when they are given a forced choice between targets and corresponding related lures (since cortex can pick up the slack in both cases). Holdstock et al. (2002) and Mayes et al. (2001) tested these predictions in a patient with focal hippocampal damage and obtained the predicted pattern of results; for additional evidence in support of these predictions, see also Westerberg et al. (2006). Contrary to this prediction, Bayley et al. (2008) observed similar levels of impairment in hippocampally lesioned patients on YN tests and forced-choice tests with corresponding related lures—however, it is possible (as is the case in all patient studies) that these patients had dysfunction in cortex that was not detected with standard MRI techniques.

To be clear, the model does not predict that all types of forced-choice responding will be relatively spared in patients with focal hippocampal damage. The prediction of relatively spared performance is related to a particular type of forced-choice test (i.e., tests that pair studied items with corresponding lures: study A, B; test A vs. A', B vs. B'). If we set up a forced-choice test that pairs studied items with lures that are related to other studied items (e.g., study A, B; test A vs. B'), the model predicts that performance will be impaired after focal hippocampal damage. This prediction has yet to be tested in patients [see Migo et al. (2009) for evidence from normal subjects that speaks to this prediction].

### **Critical Difference 3: The Cortical Signal Tracks Global Match; The Hippocampal Signal Does Not**

Numerous behavioral studies have demonstrated that subjects' recognition judgments are sensitive to global match, where global match is defined as the summed similarity of the test probe to all of the studied items (Clark and Gronlund, 1996). Studies have found that the probability of calling a non-studied test item “old” increases as a function of (1) the number of similar items that have been studied [e.g., Koutstaal et al. (1999), Shiffrin et al. (1995), Zaki and Nosofsky (2001), and Criss and Shiffrin (2004)] and (2) the degree of similarity between studied items and the test item, holding the number of studied items constant [e.g., Kahana and Sekuler (2002) and Yotsumoto et al. (2007)].

The memory signal generated by the cortical network tracks, in a graded fashion, the global match between the test cue and

all stored memory traces. This property arises because cortex assigns similar representations to similar stimuli—if the representation of one item is sharpened, this benefit accrues to other stimuli in proportion to their similarity to the initial stimulus. Importantly, the hippocampal model does not exhibit this global match property. Because of hippocampal pattern separation, the test probe has to be very similar to a studied item before it triggers any recall. If the test probe is a studied item, adding other items to the study list that overlap moderately with that item will reliably boost the cortical memory signal, but the hippocampal signal will not increase, and (depending on how the model is parameterized) the hippocampal signal can even decrease due to interference between all of the similar memory traces. This is a manifestation of the well-known principle that increasing the number of associates to a given cue can hurt recall of these associates; this principle has been described elsewhere in terms of *cue overload effects* (Watkins and Watkins, 1975) and *fan effects* (Anderson and Reder, 1999; Reder et al., 2000).

### Boundary Conditions on the Critical Differences Listed Above

The three critical properties of the hippocampal model listed above (non-Gaussian memory strength distributions leading to a positive ROC  $Y$ -intercept; good discrimination between studied items and related lures; lack of sensitivity to global match) are straightforward consequences of the hippocampal model's ability to assign relatively distinct representations to overlapping inputs. Importantly, there are limits on this ability—if the inputs coming into the hippocampus are extremely similar to one another, the hippocampus will fail to assign distinct representations to these patterns. In this “pattern separation failure” regime, the properties of the hippocampal memory signal come to resemble the properties of the cortical memory signal—the memory strength distributions for studied items and lures start to become more Gaussian and overlapping, and (correspondingly) the ROC becomes more symmetrical and the  $Y$ -intercept approaches zero. This difference in operating characteristics can be seen by comparing Figure 5 (which shows the distribution of the hippocampal memory signal given 40.5% feature overlap between input patterns) and Figure 4 (which shows the distribution of the hippocampal memory signal given 20% feature overlap between input patterns). Elfman et al. (2008) tested the model's predictions about effects of feature overlap using a source memory experiment where each item was presented either on the left or the right side of the screen at study. They manipulated the level of item feature overlap by using photos of houses as stimuli (in the “high feature overlap” condition) or else using randomly selected photos as stimuli (in the “low feature overlap” condition). At test, subjects were shown photos and asked to say whether each photo was presented on the left or right side of the screen at study. Elfman et al. (2008) plotted normalized source memory ROC curves (under conditions of high vs. low item feature overlap) and measured the shape of the z-ROC—if the underlying distributions are Gaussian, then

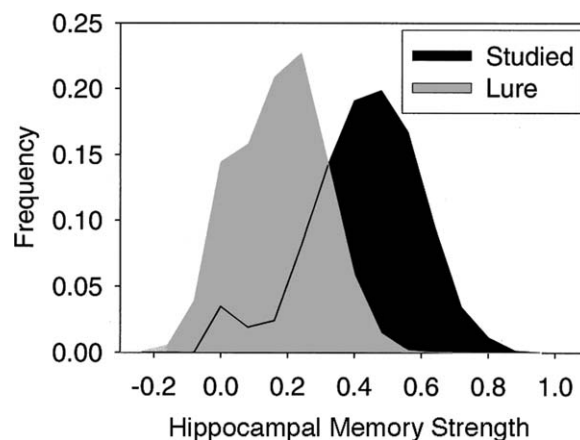


FIGURE 5. Histogram of the memory signal generated by the CLS hippocampal network, given 40.5% average overlap between input patterns. Figure adapted from Norman and O'Reilly (2003).

the z-ROC will be linear; conversely, curvature in the z-ROC implies that the underlying distributions are not Gaussian [but see Malmberg and Xu (2006) for some cautions regarding how to interpret z-ROC curvature]. In keeping with the CLS model's predictions, Elfman et al. (2008) found that the z-ROC was significantly U-shaped (indicating non-Gaussian distributions) in the low feature overlap condition and that the z-ROC was significantly less U-shaped in the high feature overlap condition, suggesting that the underlying distributions in this condition were closer to being Gaussian.

## IMPLICATIONS FOR THE CURRENT DEBATE

### Characterizing Hippocampal Versus Cortical Contributions

#### *Recollection versus familiarity*

The question of how well recollection/familiarity map onto the hippocampus/cortex distinction depends critically on how we define these terms. If we define familiarity very broadly as “a scalar signal that can be used to support recognition memory judgments,” then our simulations suggest that hippocampus can support familiarity—the average level of activity in region CA1 of the hippocampal model discriminates between old and new items, and this signal could conceivably be used to support recognition memory judgments. However, within the memory modeling community, practically all researchers have adopted a more specific definition of familiarity centered around the idea of summed similarity or global match [e.g., Gillund and Shiffrin (1984), Hintzman (1988), Nosofsky (1988), Humphreys et al. (1989), Murdock (1993), and Shiffrin and Steyvers (1997); for reviews, see Clark and Gronlund (1996), Malmberg (2008), and Norman et al. (2008)].

Our simulations show that the cortical model can support global match judgments, but the hippocampal model cannot support these judgments, except under conditions where stimuli

are extremely similar to one another. Some of the behavioral studies that have demonstrated sensitivity to global match have used extremely similar stimuli [e.g., Kahana and Sekuler (2002) used grayscale textures generated by superimposing horizontal and vertical gratings, and Yotsumoto et al. (2007) used highly similar synthetically generated faces], but global match effects can be observed even when stimuli are not highly similar to one another. For example, Shiffrin et al. (1995) purposefully disguised their main manipulation (of the number of items in a semantic category) by using loosely defined categories and intermixing items from different categories in the study list; subjects did not notice the category structure of the list, but they still showed a robust effect of category length on false alarms to items from that category. If we follow the lead of the memory modeling community and treat “sensitivity to global match” as an essential property of familiarity, our CLS model simulations suggest that cortex can support familiarity but hippocampus can not support familiarity (except in the aforementioned case where the stimuli being used in the experiment are all extremely similar to one another).

### *Moving beyond recollection versus familiarity*

It is important to not let definitional questions about recollection and familiarity distract us from the fundamental question at hand: coming up with a precise, testable description of the essential differences between the two systems’ contributions to recognition memory. Generating these precise computational descriptions was the central goal of the Norman and O’Reilly (2003) *Psychological Review* paper. However, in retrospect, it appears that we muddied the waters by continuing to use the terms “familiarity” and “recollection” (with all of their baggage) to describe cortical and hippocampal contributions. For many readers, the take-home message of our work was that it was basically OK to keep using the existing recollection/familiarity framework. Now that we are revisiting these issues, our goal is to refocus discussion on the most fundamental computational idea to come out of our simulations, namely: *Hippocampal pattern separation imbues the hippocampal signal with different operating characteristics than the cortical signal*. Pattern separation makes the hippocampus especially well-suited for discriminating between studied items and related lures; it makes the hippocampus especially poorly suited for computing global match; and it imbues the hippocampal ROC curve with a  $Y$ -intercept  $> 0$ .

The basic idea that hippocampus does pattern separation is widely accepted in the field, and the past several years have seen the publication of some extremely clever and powerful experimental demonstrations of hippocampal pattern separation [e.g., Bakker et al. (2008), Leutgeb et al. (2007), and Clelland et al. (2009)]. However, we think that the implications of pattern-separation differences (in hippocampus vs. cortex) for recognition performance have not been fully explored. For example, although there have been fMRI studies demonstrating that hippocampus does an especially good job of separating out studied items from similar lures (Bakker et al., 2008), there have not (to our knowledge) been any imaging studies that test

the model’s predictions about the respective contributions of cortex versus hippocampus to global-matching effects. These predictions could easily be tested using fMRI or intracranial EEG. Specifically, one could run experiments that parametrically manipulate the number of studied items related to each test probe [as per Shiffrin et al. (1995)] and also the similarity of these items to the test probe [as per Kahana and Sekuler (2002) and Yotsumoto et al. (2007)]. The key prediction is that, in every situation where subjects show a behavioral global match effect, perirhinal cortex activity should monotonically track global match. By contrast, hippocampal activity should only track global match in situations where all of the items are very similar to one another. If we can establish that cortex is more useful for global match computations than the hippocampus, this would be a huge step forward in fractionating the contributions of MTL substructures to recognition memory. Although the claim that “hippocampus does not implement familiarity” is subject to interpretive ambiguity relating to the meaning of “familiarity,” the claim that “hippocampus cannot compute summed similarity (unless the inputs show extensive overlap)” is much more clear.

### *Implications for measurement models*

The CLS model simulations described earlier can help us to assess the validity of techniques that are commonly used to measure the contributions of hippocampus and cortex to recognition memory. Here, we will focus on the dual-process signal-detection (DPSD) model set forth by Yonelinas et al. (Yonelinas, 1994, 2001; Yonelinas et al., 1998, 2002), because this is the most widely used measurement model in the literature. The most important assumptions underlying the DPSD measurement model are (1) that familiarity is an equal-variance Gaussian signal detection process (yielding a symmetric ROC with a zero  $Y$ -intercept) and (2) recollection is a thresholded process (in the technical sense that the studied-item evidence distribution extends to the right of the lure evidence distribution, yielding an asymmetric ROC with a positive  $Y$ -intercept). Given these assumptions, any asymmetry in the ROC curve can be attributed to the recollection process. If we make the further assumption that recollection maps onto the hippocampus (Yonelinas, 2002), we can use asymmetry in the ROC to index the degree to which the hippocampus is contributing to recognition performance.

According to the CLS model, the assumptions underlying the DPSD model will hold in some circumstances but not others. In our simulations, the cortical signal is well-described by an equal-variance Gaussian signal-detection model unless contribution of encoding variability (which affects targets more than lures) is large relative to the contribution of other forms of variability (e.g., variability in pre-experimental exposure to stimuli) that affect targets and lures equally. Also, in our simulations, the hippocampal signal yields a positive ROC  $Y$ -intercept unless there is an extremely high level of feature overlap between items. If encoding variability is controlled, for example, using an encoding task, and between-item similarity is not too high, then esti-



mates of recollection and familiarity obtained using the DPSD procedure should approximately reflect the contributions of the hippocampal and cortical processes (respectively). More specifically, the DPSD estimate of hippocampal recollection corresponds to the height of the (estimated) ROC  $Y$ -intercept, which indicates the proportion of the studied item distribution that lies to the right of the maximum of the lure distribution. The hippocampal distributions shown in Figure 4 are too complex to be fully described with this single number—knowing the proportion of the studied-item distribution that lies above a particular memory strength value does not tell us how the distribution is shaped below that point. Nonetheless, the  $Y$ -intercept is a useful proxy for the overall strength of the hippocampal signal in the model.

Our CLS simulations suggest that it should not be difficult to engineer situations where the assumptions of DPSD are invalid. For example, strong encoding variability will lead to violations of the equal-variance assumption for the cortical process [strong encoding variability will also lead to violations of the assumption that the two signals are stochastically independent; see and Norman and O'Reilly (2003) for further discussion of this point]. It seems clear to us that blindly using models like DPSD without regard to boundary conditions can lead to trouble. However, while the assumptions of DPSD will not always be met, it seems likely to us that many—if not most—of the studies that have used DPSD fit within the boundary conditions outlined earlier (low-encoding variability, not too much overlap between items in the learning set). For example, in the studies that have applied ROC analysis to data from hippocampally lesioned animals (Fortin et al., 2004; Sauvage et al., 2008), the rats had to sniff each odor at training (thereby reducing encoding variability), and the odors were very distinctive (thereby ensuring that there would not be too much overlap between the cortical representations of these odors).

Speaking generally, a measurement model does not have to be exactly correct to be a useful source of converging evidence and it does not have to be applicable in every situation. For these reasons, it seems like we would be depriving ourselves of a useful tool if we banished measurement models like DPSD. We just need to be aware of the boundary conditions, and mechanistic models like CLS give us a way of developing intuitions about what these boundary conditions will be.

## SUMMARY

The high-level goal of this work has been to reassert the importance of mechanistic models in the debate over how hippocampus and cortex contribute to recognition memory. It is worth noting that practically all the recent debates over the neural basis of recognition memory have focused on the relative validity of different measurement models—the DPSD model versus the unequal-variance signal-detection model described by Wixted (2007) and others versus the Shimamura and Wickens (2009) model versus countless other models [for a review, see

Yonelinas and Parks (2007)]. Mechanistic models like CLS have played a relatively small role in the debate thus far. In this work, we have described two related areas where mechanistic models can contribute: First, mechanistic models can be used to proactively generate predictions about how particular manipulations will affect hippocampal and cortical memory processes. Second, mechanistic models can be used to evaluate the assumptions of measurement models.

The specific goal of this work has been to reassert the claim from Norman and O'Reilly (2003) that pattern-separation differences are the key to understanding differences in hippocampal versus cortical differences. We listed several implications of these pattern-separation differences for recognition performance, including the idea that cortex is better-suited for computing global match. We believe that fully exploring these implications, using models like CLS as a guide, will help us to arrive at a richer understanding of how the contributions of hippocampus versus cortex to recognition memory differ.

## REFERENCES

- Aggleton JP, Brown MW. 1999. Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behav Brain Sci* 22:425–490.
- Anderson JR, Reder LM. 1999. The fan effect: New results and new theories. *J Exp Psychol: Gen* 128:186.
- Bakker A, Kirwan CB, Miller M, Stark CEL. 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science (New York, NY)* 319:1640–1642.
- Barense MD, Bussey TJ, Lee AC, Rogers TT, Davies RR, Saksida LM, Murray EA, Graham KS. 2005. Functional specialization in the human medial temporal lobe. *J Neurosci* 25:10239–10246.
- Bayley PJ, Wixted JT, Hopkins RO, Squire LR. 2008. Yes/no recognition, forced-choice recognition, and the human hippocampus. *J Cogn Neurosci* 20:505–512.
- Becker S. 2005. A computational principle for hippocampal learning and neurogenesis. *Hippocampus* 15:722–738.
- Bogacz R, Brown MW. 2003. Comparison of computational models of familiarity discrimination in the perirhinal cortex. *Hippocampus* 13:494–524.
- Burgess N, O'Keefe J. 1996. Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus* 6:749–762.
- Bussey TJ, Saksida LM. 2002. The organisation of visual object representations: A connectionist model of effects of lesions in perirhinal cortex. *Eur J Neurosci* 15:355–364.
- Bussey TJ, Saksida LM, Murray EA. 2002. The role of perirhinal cortex in memory and perception: Conjunctive representations for object identification. In: Witter MP, Waterlood FG, editors. *The Parahippocampal Region: Organisation and Role in Cognitive Functions*. New York: Oxford.
- Clark SE, Gronlund SD. 1996. Global matching models of recognition memory: How the models match the data. *Psychonom Bull Rev* 3, 37–60.
- Clelland CD, Choi M, Romberg C, Clemenson GD, Fagniere A, Tyers P, Jessberger S, Saksida LM, Barker RA, Gage FH, Bussey TJ. 2009. A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science (New York, NY)* 325:210–213.
- Cowell RA, Bussey TJ, Saksida LM. 2006. Why does brain damage impair memory? A connectionist model of object recognition memory in perirhinal cortex. *J Neurosci* 26:12186–12197.

- Cowell RA, Bussey TJ, Saksida LM. 2010. Functional dissociations within the ventral object processing pathway: Cognitive modules or a hierarchical continuum? *J Cogn Neurosci* 22:2460–2479.
- Cowell RA, Bussey TJ, Saksida LM. (this issue). Components of recognition memory. Dissociable cognitive processes or just differences in representational complexity? *Hippocampus*.
- Criss AH, Shiffrin RM. 2004. Context noise and item noise jointly determine recognition memory: A comment on dennis and humphreys (2001). *Psychol Rev* 111:800–807.
- Diana RA, Yonelinas AP, Ranganath C. 2007. Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends Cogn Sci* 11:379–386.
- Eichenbaum H, Otto T, Cohen NJ. 1994. Two functional components of the hippocampal memory system. *Behav Brain Sci* 17:449–518.
- Elfman KW, Parks CM, Yonelinas AP. 2008. Testing a neurocomputational model of recollection, familiarity, and source recognition. *J Exp Psychol Learn Mem Cogn* 34:752–768.
- Fortin NJ, Wright SP, Eichenbaum HB. 2004. Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature* 431:188–191.
- Gardiner JM, Brandt KR, Baddeley AD, Vargha-Khadem F, Mishkin M. 2008. Charting the acquisition of semantic knowledge in a case of developmental amnesia. *Neuropsychologia* 46:2865–2868.
- Gillund G, Shiffrin RM. 1984. A retrieval model for both recognition and recall. *Psychol Rev* 91:1–67.
- Grossberg S. 1976. Adaptive pattern classification and universal coding. I. Parallel development and coding of neural feature detectors. *Biol Cybernet* 23:121–134.
- Grossberg S. 1986. The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In: Swab EC, Nusbaum HC, editors. *Pattern Recognition in Humans and Machines, Vol. 1: Speech Perception*. New York: Academic Press. p 187–294.
- Grossberg S, Stone G. 1986. Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychol Rev* 93:46–74.
- Hasselmo ME, Wyble B. 1997. Free recall and recognition in a network model of the hippocampus: Simulating effects of scopolamine on human memory function. *Behav Brain Res* 89:1–34.
- Hilford A, Glanzer M, Kim K, Decarlo LT. 2002. Regularities of source recognition: ROC analysis. *J Exp Psychol: Gen* 131:494–510.
- Hintzman DL. 1988. Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychol Rev* 95:528–551.
- Hintzman DL, Curran T, Oppy B. 1992. Effects of similarity and repetition on memory: Registration without learning. *J Exp Psychol Learn Mem Cogn* 18:667–680.
- Holdstock JS, Mayes AR, Roberts N, Cezayirli E, Isaac CL, O'Reilly RC, Norman KA. 2002. Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus* 12:341–351.
- Humphreys MS, Bain JD, Pike R. 1989. Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychol Rev* 96:208–233.
- Jacoby LL, Yonelinas AP, Jennings JM. 1997. The relation between conscious and unconscious (automatic) influences: A declaration of independence. In: Cohen JD, Schooler JW, editors. *Scientific Approaches to Consciousness*. Mahwah, NJ: Erlbaum. pp 13–47.
- Kahana MJ, Sekuler R. 2002. Recognizing spatial patterns: A noisy exemplar approach. *Vis Res* 42:2177–2192.
- Koutstaal W, Schacter DL, Jackson EM. 1999. Perceptually based false recognition of novel objects in amnesia: Effects of category size and similarity to category prototypes. *Cogn Neuropsychol* 16:317.
- Leutgeb JK, Leutgeb S, Moser M-B, Moser EI. 2007. Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science (New York, NY)* 315:961–966.
- Li L, Miller EK, Desimone R. 1993. The representation of stimulus familiarity in anterior inferior temporal cortex. *J Neurophysiol* 69:1918–1929.
- Macmillan NA, Creelman CD. 2005. *Detection Theory: A User's Guide*, 2nd ed. Mahwah, NJ: Erlbaum.
- Malmberg KJ. 2008. Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cogn Psychol* 57:335–384.
- Malmberg KJ, Xu J. 2006. The influence of averaging and noisy decision strategies on the recognition memory ROC. *Psychonom Bull Rev* 13:99–105.
- Marr D. 1971. Simple memory: A theory for archicortex. *Philos Trans R Soc (Lond) B* 262:23–81.
- Mayes AR, Isaac CL, Downes JJ, Holdstock JS, Hunkin NM, Montaldi D, MacDonald C, Cezayirli E, Roberts JN. 2001. Memory for single items, word pairs, and temporal order in a patient with selective hippocampal lesions. *Cogn Neuropsychol* 18:97–123.
- McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102:419–457.
- McNaughton BL, Morris RGM. 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci* 10:408–415.
- Migo E, Montaldi D, Norman KA, Quamme J, Mayes A. 2009. The contribution of familiarity to recognition memory is a function of test format when using similar foils. *Quart J Exp Psychol* 62:1198–1215.
- Moll M, Miikkulainen R. 1997. Convergence-zone episodic memory: Analysis and simulations. *Neural Networks* 10:1017–1036.
- Murdock BB. 1993. TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychol Rev* 100:183–203.
- Norman KA, O'Reilly RC. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychol Rev* 104:611–646.
- Norman KA, Newman EL, Perotte AJ. 2005. Methods for reducing interference in the complementary learning systems model: Oscillating inhibition and autonomous memory rehearsal. *Neural Networks* 18:1212–1228.
- Norman KA, Newman EL, Detre GJ, Polyn SM. 2006. How inhibitory oscillations can train neural networks and punish competitors. *Neural Comput* 18:1577–1610.
- Norman KA, Detre GJ, Polyn SM. 2008. Computational models of episodic memory. In: Sun R, editor. *The Cambridge Handbook of Computational Psychology*. New York: Cambridge University Press.
- Nosofsky R. 1988. Exemplar-based accounts of relations between classification, recognition, and typicality. *J Exp Psychol: Learn Mem Cogn* 14:700708.
- O'Keefe J, Nadel L. 1978. *The Hippocampus as a Cognitive Map*. Oxford, UK: Oxford University Press.
- O'Reilly RC, Rudy JW. 2001. Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychol Rev* 108:311–345.
- O'Reilly RC, Norman KA. 2002. Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends Cogn Sci* 6:505–510.
- Ranganath, C. (this issue). A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus*.
- Reder LM, Nhouyvanisvong A, Schunn CD, Ayers MS, Angstadt P, Hiraki KA. 2000. A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *J Exp Psychol: Learn Mem Cogn* 26:294–320.
- Rolls ET. 1989. Functions of neuronal networks in the hippocampus and neocortex in memory. In: Byrne JH, Berry WO, editors.

- Neural Models of Plasticity: Experimental and Theoretical Approaches. San Diego, CA: Academic Press. pp 240–265.
- Sauvage MM, Fortin NJ, Owens CB, Yonelinas AP, Eichenbaum H. 2008. Recognition memory: Opposite effects of hippocampal damage on recollection and familiarity. *Nat Neurosci* 11:16–18.
- Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21.
- Sherry DF, Schacter DL. 1987. The evolution of multiple memory systems. *Psychol Rev* 94:439–454.
- Shiffrin RM, Huber DE, Marinelli K. 1995. Effects of category length and strength on familiarity in recognition. *J Exp Psychol: Learn Mem Cogn* 21:267–287.
- Shiffrin RM, Steyvers M. 1997. A model for recognition memory: REM—retrieving effectively from memory. *Psychonom Bull Rev* 4:145–166.
- Shimamura AP, Wickens TD. 2009. Superadditive memory strength for item and source recognition: The role of hierarchical relational binding in the medial temporal lobe. *Psychol Rev* 116:1–19.
- Squire LR. 1992. Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99:195–231.
- Squire LR, Wixted JT, Clark RE. 2007. Recognition memory and the medial temporal lobe: A new perspective. *Nature reviews. Neuroscience* 8:872–883.
- Sutherland RJ, Rudy JW. 1989. Configural association theory: The role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology* 17:129–144.
- Teyler TJ, Discenna P. 1986. The hippocampal memory indexing theory. *Behav Neurosci* 100:147–154.
- Treves A, Rolls ET. 1994. A computational analysis of the role of the hippocampus in memory. *Hippocampus* 4:374–392.
- Viskontas IV, Knowlton BJ, Steinmetz PN, Fried I. 2006. Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *J Cogn Neurosci* 18:1654–1662.
- Watkins OC, Watkins MJ. 1975. Buildup of proactive inhibition as a cue-overload effect. *Hum Learn* 104:442–452.
- Westerberg CE, Paller KA, Weintraub S, Mesulam MM, Holdstock JS, Mayes AR, Reber PJ. 2006. When memory does not fail: Familiarity-based recognition in mild cognitive impairment and Alzheimer's disease. *Neuropsychology* 20:193–205.
- Wixted JT. 2007. Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev* 114:152–176.
- Wixted JT, Squire LR. 2008. Constructing receiver operating characteristics (ROCs) with experimental animals: Cautionary notes. *Learn Mem* 15:687–690.
- Wu X, Baxter RA, Levy WB. 1996. Context codes and the effect of noisy learning on a simplified hippocampal CA3 model. *Biol Cybernet* 74:159–165.
- Xiang JZ, Brown MW. 1998. Differential encoding of novelty, familiarity, and recency in regions of the anterior temporal lobe. *Neuropharmacology* 37:657–676.
- Yonelinas AP. 1994. Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *J Exp Psychol: Learn Mem Cogn* 20:1341–1354.
- Yonelinas AP. 2001. Consciousness, control, and confidence: The 3 Cs of recognition memory. *J Exp Psychol: Gen* 130:361–379.
- Yonelinas AP. 2002. The nature of recollection and familiarity: A review of 30 years of research. *J Mem Lang* 46:441–517.
- Yonelinas AP, Parks CM. 2007. Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychol Bull* 133:800–832.
- Yonelinas AP, Kroll NEA, Dobbins I, Lazzara M, Knight RT. 1998. Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology* 12:323–339.
- Yonelinas AP, Kroll NE, Quamme JR, Lazzara MM, Sauve MJ, Widaman KF, Knight RT. 2002. Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nat Neurosci* 5:1236–1241.
- Yotsumoto Y, Kahana MJ, Wilson HR, Sekuler R. 2007. Recognition memory for realistic synthetic faces. *Mem Cogn* 35:1233–1244.
- Zaki S, Nosofsky R. 2001. Exemplar accounts of blending and distinctiveness effects in perceptual old-new recognition. *J Exp Psychol: Learn Mem Cogn* 27:10221041.