Neural evidence of recognition—but not recall—for videos in monkey hippocampus

John J. Sakon[a] & Wendy A. Suzuki[b]

[a]Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104
[b]Center for Neural Science, New York University, New York, NY, 10003
Correspondence: sakon@upenn.edu

Abstract

The role of the hippocampus in recognition memory vs. recall has long been a source of debate. Tasks used to study recognition memory that almost exclusively require an explicit probe—where the subject must make a response to prove they remember—yield mixed results on hippocampal involvement. Meanwhile, almost all tasks probing hippocampal recall also use explicit probes. Here, instead of requiring such an artificial probe, we tasked monkeys to freely view naturalistic videos, and only tested their memory implicitly via a novel v. repeat video condition. Notably, a large proportion of hippocampal neurons differentiated novel from repeat videos via firing rates, supporting the involvement of the hippocampus in implicit recognition memory. The majority of these neurons showed an enhancement in firing rate to repeat instead of novel videos, a pattern which has not previously been shown in hippocampus. However, there was no evidence of post-stimulus cued recall, as neurons only showed these signals during video stimulus presentation. These results suggest the hippocampus is involved in recognition memory for dynamic visual stimuli when memory retrieval is not explicitly probed.

Introduction

The ability to identify a previously experienced stimulus is termed recognition memory(Brown and Xiang, 1998). There is debate as to whether the hippocampus—a brain region typically linked to associative and episodic memory(Eichenbaum et al., 2012)—contributes to recognition memory(Bird, 2017). Lesion studies in monkeys and amnesic studies in humans suggest that the hippocampus may only be involved in recognition memory depending on task parameters like stimulus type and frequency of stimulus repetition during and across sessions (Baxter, 2009; Bird, 2017). Neurophysiological studies using both human fMRI and monkey electrophysiology have typically found that regions in and adjacent to the medial temporal lobe such as entorhinal cortex, perirhinal cortex, and inferotemporal cortex show evidence of recognition memory via decreases in responses to repeated stimuli (i.e. repetition suppression(Barron et al., 2016; Meyer and Rust, 2018)), while electrophysiological studies of hippocampus have typically reported little or no such neural signals(Brown et al., 1987; Brown and Xiang, 1998; Riches et al., 1991; Rolls et al., 1993; Xiang and Brown, 1998).

However, recent studies in monkeys have shown electrophysiological evidence of recognition memory in the hippocampus, with separate subgroups of neurons either increasing or decreasing their firing rates in response to presentation of repeated stimuli(Jutras and Buffalo, 2010; Sakon and Suzuki, 2019). A key distinction in the tasks that find recognition memory signals in hippocampus is that animals were not required to make an explicit response to probe their memory as had been required in previous work(Brown et al., 1987; Riches et al., 1991; Rolls et al., 1993; Xiang and Brown, 1998). While the distinction between implicit and explicit tasks(Voss and Paller, 2008) might seem slight, studies in humans have found that tasks requiring explicit memory probes—instead of incidental probes(Bakker et al., 2008; Lacy et al., 2011)—alter the neural responses in hippocampus, possibly because the requirement of a decision recruits additional brain regions that support strategy (e.g. recall-to-reject(Lacy et al., 2011)) or integration of evidence from memory(Shadlen and Shohamy, 2016).

In humans, the hippocampus is typically associated with recollection(Baldassano et al., 2017; Chadwick et al., 2010; Dimsdale-Zucker et al., 2018; Horner et al., 2015), and work in humans suggests the hippocampus can contribute to recognition memory via both recollection and familiarity (Merkow et al., 2015; Rutishauser et al., 2008; Wais et al., 2010; Wais et al., 2006). However, electrophysiological studies of human recognition memory almost exclusively use task designs with explicit memory probes (old/new judgments)(Fried et al., 1997; Heit et al., 1990; Merkow et al., 2015; Rutishauser et al., 2008; Viskontas et al., 2006), even though the familiarity component of recognition memory is more closely tied to implicit memory processes(Yonelinas, 2002). One notable exception is a study(Gelbard-Sagiv et al., 2008) of epileptic patients with surgically implanted electrodes that screened for hippocampal neurons that increased in firing rate to specific cinematic episodes (e.g. The Simpsons). These same neurons then responded in a similar

manner when the patients freely recalled the episode they had recently viewed. To the best of our knowledge, neural recording in monkeys freely viewing such video stimuli has not been performed, making it unknown if the electrophysiological signatures of implicit recognition memory shown in monkey hippocampus detailed above (Jutras and Buffalo, 2010; Sakon and Suzuki, 2019) hold true when animals recall memories of such episodic stimuli.

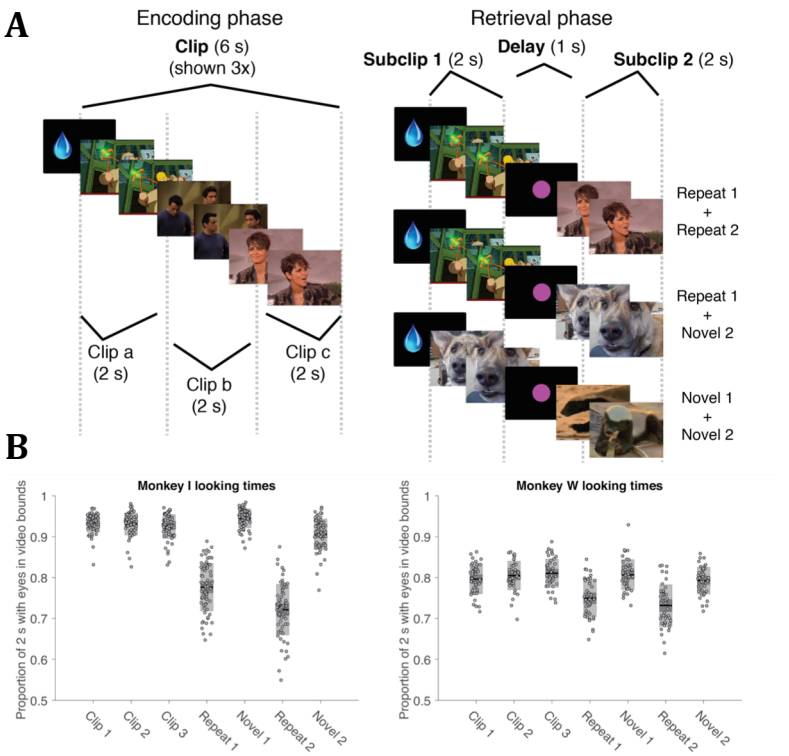


Figure 1. Task design and behavior. **A)** In each trial, monkeys were shown 6 s clips, which were comprised of three, 2 s movies (sections a, b, and c) spliced together from random video sources (Methods). Each clip was repeated 3x during the encoding phase. In the retrieval phase, monkeys were shown two, 2 s subclips surrounding a 1 s delay (a delay of only 1 s was used to maintain monkey attention). Three trial types were equally likely during this phase: i) subclip 1 is a repeat of the first 2 s of the clips and subclip 2 is a repeat of the last 2 s of the clips, ii) subclip 1 is a repeat of the first 2 s of the clips and subclip 2 is a novel video, and iii) subclip 1 and subclip 2 are both novel videos. Each liquid drop icon indicates a fixation "game" where animals were given juice reward for achieving 0.5 s fixation to a dot. After reward was given, monkeys had to re-acquire fixation for 0.5 s to begin the upcoming video. **B)** The proportion of time each monkey's eye gazes were within the bounds of the movie for section a of clips 1-3, and for each of the 2 s subclip periods, averaged over sessions. Note that Repeat 1 (shown during subclip 1) is a repeat of the first 2 s of the clips (section a), while Repeat 2 (shown during subclip 2) is a repeat of the last 2 s of the clips (section c). Therefore, the middle 2 s of the clip (section b) is never shown during retrieval. Shaded boxes are 95% confidence intervals (dark gray) and the SDs (light gray).

Therefore, to separately study hippocampal responses to both implicit recognition and recall, we recorded single neurons from monkey hippocampus as they engaged in a task with both a recognition condition and a distinct cued recall period (Fig. 1A). Monkeys freely viewed 6 s episodes ("**clips**"), comprised of three, contiguous 2 s videos (each 2s video designated a, b, and c), sampled from Internet videos of people, animals, animations, and other such engaging content. In each trial, after three presentations of the 6 s clips (the encoding phase), the animals were then shown two, 2 s "**subclips**" with a 1 s delay period in between (the retrieval phase). In subclip 1, either the first 2 s of the original clip was shown (repeat 1), or a new 2 s subclip was shown (novel 1). Following the 1 sec delay period, subclip 2 could be either the last

2 s of the original clip (repeat 2), or a new 2 s subclip (novel 2). Crucially, once a clip or subclip was used in a given trial, it was never again used for that monkey, thereby making the novel videos a true first presentation, and the repeat videos targeted probes of memory for episodes(Ezzyat and Davachi, 2011).

This task design gives us two key periods to separably investigate recognition and recall. First, the novel v. repeat conditions in the two subclip periods allow us to probe recognition of the recently seen clips. Since we are using an implicit, free-viewing paradigm, we anticipate hippocampal neurons will show evidence of recognition memory via changes in firing rate to repeat v. novel videos similar to during serial image tasks(Jutras and Buffalo, 2010; Sakon and Suzuki, 2019). Second, the delay period between the subclips allows us to investigate cued recall of the clip period. Specifically, when subclip 1 is a repeat of the first 2 s of the clip period, we anticipate this cue will stimulate recall of the remainder of the clip during the delay, and this cued recall will be reflected in our population of hippocampal neurons.

Results

Behavior. To determine if monkeys engaged with the task, we looked for evidence of each animal remembering videos during each trial via a preferential looking paradigm, in which memory is associated with viewing repeated stimuli for shorter times than novel stimuli(Jutras and Buffalo, 2010; Sakon and Suzuki, 2019; Spelke, 1985). We measured the time each monkey looked within the bounds of the video stimuli during the first 2 s of all three clips, as well as during the two, 2 s subclip periods for 110 sessions (Fig. 1A; 71 monkey I and 39 monkey W). The looking times for each of these periods are averaged for each session and then the proportion of the 2 s the animal looked at the video is plotted separately for each monkey (Fig. 1B). We anticipate that for both subclip 1 and subclip 2, when monkeys are shown a repeat instead of a novel video, on average they should look for less time than novels (note that, when either subclip is a repeat, it is the 4th presentation of that video, since all repeat videos were also shown during clips 1-3). Indeed, for each monkey, both during the subclip 1 (Monkey I, p=5.8e-24; Monkey W, p=2.4e-7; Wilcoxon rank-sum test, FWE-corrected) and the subclip 2 (Monkey I, p=2.1e-23; Monkey W, p=2.4e-7; Wilcoxon rank-sum test) periods, monkeys looked at repeat videos for significantly less time than novel videos. Each monkey also looked at the novel videos during the subclip 1 period as long or longer than the first presentation of the clip (Monkey I, 0.933 for clip 1 v. 0.945 for subclip 1, p=0.0049; Monkey W, 0.807 for clip 1 v. 0.797 for subclip 1, p=1.0; Wilcoxon rank-sum test, FWE-corrected). The significant differences in looking times between novel and repeat videos, in addition to the high proportion of time monkeys looked at novel videos (>0.90 for monkey I and >0.79 for monkey W for all novel clips and subclips), indicate monkeys were engaged in the task.

To quantify how well monkeys classify repeat v. novel videos in our task, we used looking times in a signal detection theory framework similar to a recent

publication(Sakon and Suzuki, 2019). With this framework, for each trial we compare the looking time between novel/repeat videos from the 1st subclip v. the first 2 s of the clip 1 period (clip 1a) using an array of time thresholds (e.g. did the monkey look at subclip 1 80%/90%/100% as long as at clip 1) to populate an ROC curve (Methods). Then, we take the threshold that maximizes d' for each session, meaning it maximizes when the monkey treats a novel subclip as novel (i.e., true positive) while also minimizing when the monkey treats a repeat subclip as novel (i.e., false positive). When we apply this process to the 110 sessions across both monkeys (average of 125.9±2.5 trials per session), on average for the 1st subclip we can classify 80.9±1.6% of novel videos as novel compared to 54.9±1.6% of repeat videos as novel, which yields d'=0.88±0.042 (SE for all errors). In addition, when we use the same process to compare the looking times from the novel/repeat videos from the 2nd subclip v. the last 2 s of the clip 1 period (clip 1c), we can classify 75.4±1.5% of novel videos as novel compared to 42.8±1.4% of repeat videos as novel, which yields d'=0.96±0.039. This strong classification of the subclip periods gives us additional behavioral evidence that the monkeys can recognize videos in both subclip periods that were recently shown during the clip periods.

Single neuron evidence of recognition memory. We recorded from 209 single units in the hippocampus (Fig. S1) of the same two monkeys from our behavioral analysis during 110 sessions (unit breakdown: Rhesus I = 120 units, Rhesus W = 89 units). In the limited previous work that has used naturalistic video stimuli during electrophysiological or neurophysiological recordings—all in humans—neurons are responsive throughout the medial temporal lobe(Gelbard-Sagiv et al., 2008) and hippocampus(Ben-Yakov et al., 2013; Chadwick et al., 2010). Therefore, while we limited our recordings to hippocampus due to its essential role in episodic memory(Eichenbaum et al., 2012), we did not localize our recordings to a particular hippocampal subfield but grouped our neurons into a single pool (note that during such deep brain acute recordings localization is limited to ±1 mm, which would make the subfield within hippocampus for many units uncertain). The firing rates of the neural population during video presentation were variable [mean firing rate ± SD for 155 neurons <10 Hz = 2.8 ± 2.6 Hz; mean firing rate ± SD for 55 putative interneurons ≥10 Hz (32) = 26.6 ± 16.8 Hz] but with a similar breakdown as our previous work recording hippocampal neurons during serial image presentation(Sakon and Suzuki, 2019). The population sparseness(Rolls and Treves, 2011) of neural firing to novel and repeat videos was $a^P = 0.30$, similar but slightly more sparse than the $a^P = 0.33$(Rolls and Treves, 2011) and $a^P = 0.37$(Sakon and Suzuki, 2019) found in previous studies of monkey hippocampus.

Importantly, we made no effort to select for stimulus-responsive neurons, but instead during daily recordings lowered the electrode into hippocampus and recorded from the first stable single unit(s) we could isolate. Monkeys were head-fixed during this process and rested in front of a blank computer screen. Therefore, our hippocampal population was not pre-selected in any way for visual or video stimulus responsiveness.

First, we focused on neural responses during the subclip periods, since they contain the key memory retrieval conditions (novel v. repeat). The proportions of neurons that changed their firing rates during novel or repeat video presentation compared to fixation periods are at similar levels as previously found in monkey hippocampus during presentation of novel or repeat images, with ~50% of neurons significantly different (Table 1, Total video responsive). However, the proportion of neurons that increased in firing rates to repeat videos, with 44% increasing compared to only 20% decreasing during subclip 1 (p=6.8e-5, X^2-test, FWE-corrected), and 31% increasing compared to only 13% decreasing during subclip 2 (p=3.0e-4, X^2-test, FWE-corrected), were both larger than expected by chance. This greater number of neurons increasing rather than decreasing in firing rate to repeat videos is surprising, considering 1) that repeated stimuli tend to show repetition suppression of neural responses in medial temporal lobe(Barron et al., 2016) and 2) previous results in monkey(Jutras and Buffalo, 2010; Sakon and Suzuki, 2019) and human(Viskontas et al., 2006) hippocampus have found the opposite trend when subjects are shown repeat images, with a greater proportion of neurons showing a

Firing rate compared to baseline	Video task: subclip 1 (144 neurons)		Video task: subclip 2 (156 neurons)		Visual preferential looking task (131 neurons)		Pattern separation task (134 neurons)	
	Novel (%)	Repeat (%)	Novel (%)	Repeat (%)	Novel (%)	Repeat (%)	Novel (%)	Repeat (%)
Increase in firing rate	27	44 *	33	31 **	20	19	25	19
Decrease in firing rate	24	20	28	13	33	29	33	28
Total video responsive	51	64	55	44	53	48	58	47
					Jutras & Buffalo (2010)		*Sakon & Suzuki (2019)*	

Table 1. Comparison of stimulus responsive neurons between video task and image recognition tasks. The percentage of monkey hippocampal neurons that significantly increase or decrease their firing rates from baseline for the subclip periods of the video task (p<0.05, Wilcoxon rank-sum) as well as image presentation in two previously published recognition memory tasks(Jutras and Buffalo, 2010; Sakon and Suzuki, 2019). The bottom row is the total percentage of stimulus responsive neurons (the sum of the first two rows). Note that the delay period compares trials after a repeat or novel video is shown (same breakdown as subclip 1). *p=6.8e-5, **p=3.0e-4.

decrease than increase in firing rate in these studies (Table 1, last 2 columns).

Next, we directly compared neural responses to repeat v. novel videos. During the subclip 1 period, 27/144 (19%) neurons fired significantly more to repeat than novel videos (p<0.05, rank-sum test), as compared to only 7/144 (5%) where neurons fired more to novel than repeat videos. These differences in proportions are not expected by chance (p=1.0e-e, X^2-test, FWE-corrected). Similarly, during the subclip 2 period, 21/156 (13%) neurons fired significantly stronger to repeat than novel videos, as compared to 8/156 (5%) where neurons fired more to novel than repeat videos, which again is not expected by chance (p=0.04, X^2-test, FWE-corrected). Once again, comparing to data from hippocampal neurons recorded

while monkeys performed a serial image task in our previous work(Sakon and Suzuki, 2019), 12/134 (9%) neurons showed significantly greater firing rates to repeat than novel images, while 15/134 (11%) showed significantly greater firing rates to novel than repeat images. Therefore, unlike in the serial image task, hippocampal neurons recorded during the video task respond proportionally more to repeat than novel stimuli.

Firing rate (FR) by condition	Video task: subclip 1 (144 neurons)	Video task: subclip 2 (156 neurons)	Pattern separation task (134 neurons)
repeat FR > novel FR	27 (19%)	21 (13%)	12 (9%)
novel FR > repeat FR	7 (5%)	8 (5%)	15 (11%)
	*	‡	

Sakon & Suzuki (2019)

Table 2. Comparison of novel v. repeat responsive neurons between video task and image recognition task. The number (%) of monkey hippocampal neurons (with FR≥0.5 Hz during given period) that respond with significantly stronger firing rates ($p<0.05$, Wilcoxon rank-sum) between novel and repeat stimuli for given for each task. *p=1.0e-3, ‡p=0.04.

We made similar comparisons after using our behavioral measure to classify trials, wherein we only compare firing rates for "repeat"|repeat (when the looking time indicates the monkey recognized the video as a "repeat" when a repeat video was actually shown) v. "novel"|novel (when the looking time indicates the monkey identified the video as a "novel" when a novel video was actually shown) trials. Using these behaviorally-identified trials, during the subclip 1 period, 36/167 (22%) neurons fired significantly stronger to repeat than novel videos, as compared to only 11/167 (7%) where neurons fired more to novel than repeat videos (p=1.7e-4, X^2-test, FWE-corrected). During the subclip 2 period, 34/170 (20%) neurons fired significantly stronger to repeat than novel videos, as compared to only 3/170 (2%) where neurons fired more to novel than repeat videos (p=1.3e-7, X^2-test, FWE-corrected). Notably, of the 36 neurons with higher firing rates to repeat videos in subclip 1 and the 34 neurons with higher firing rates to repeat videos in subclip 2, 22 of these neurons are the same, indicating these neurons are tuned to repeat video detection regardless of the period.

Population evidence of recognition memory. Considering the strong, differentiable responses at the individual neuron level to novel and repeat videos, we asked if we could classify novel and repeat videos in the subclip periods via the firing rates of our neural population. First, focusing on subclip 1, we took the 110 neurons with a minimum firing rate of 0.5 Hz (69 neurons <10 Hz = 2.6 ± 2.5 Hz; 41 putative interneurons ≥10 Hz = 25.2 ± 20.1 Hz) that were recorded for at least 35 novel and 35 repeat trials and created a 10-fold, cross-validated logistic regression classifier (Kiani et al., 2014; Sakon and Suzuki, 2019) on this pseudopopulation. This classifier trains each fold on 90% of trials and tests on the held out 10% in a balanced design, thereby using every trial once in one of the ten test sets. Since we use the neural responses to predict a binary classification, we are only testing if hippocampal neurons generally categorize novel v. repeat videos agnostic to the content of the videos themselves. We trained separate classifiers for spikes integrated over 500 ms time windows that were stepped through in 250 ms

increments locked onto stimulus presentation. As most neurons in the pseudopopulation were recorded for more than 35 novel or repeat trials (mean trials ± SE: 79.8 ± 1.5 novel and 77.5 ± 1.3 repeats), we ran each classifier for 200 permutations and randomly subsampled from the total trials recorded for each neuron. The average prediction accuracy from these permutations is shown in Fig. 2, with significant clusters of classification assessed using a maximum statistic method that accounts for multiple comparisons over time(Maris and Oostenveld, 2007) (Methods).

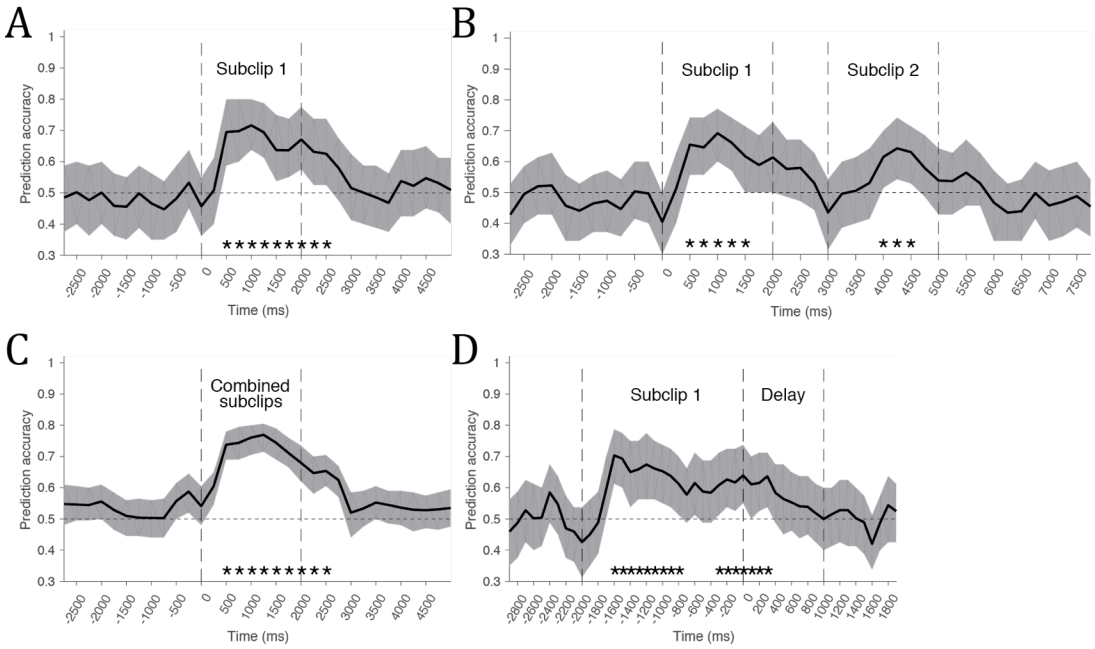


Figure 2. Classification of novel and repeat videos via firing rates of neural population. **A)** Prediction accuracy of novel v. repeat video identification aligned to subclip 1 onset for all trials combined. Dashed lines indicate subclip 1. **B)** Prediction accuracy when classifying trials where the monkey was shown either novel+novel (1/3rd of trials) or repeat+repeat (1/3rd of trials) videos for subclip 1+subclip 2. Dashed lines outline the contiguous subclip 1 (2 s), delay (1 s), and subclip 2 (2 s) periods. **C)** Prediction accuracy when classifying pooled subclip 1 and subclip 2 trials based on whether novel or repeat videos were shown during that period. **D)** Prediction accuracy aligned to delay onset with trials classified on whether novel or repeat videos were shown during preceding subclip 1 period. Spikes are integrated in windows of 500 ms with a moving window of 250 ms for each point in the first 3 graphs, and integrated for 200 ms with 100 ms steps for the last graph to show higher resolution of the drop-off after delay onset. Error bars for all graphs are the 5th and 95th percentile of prediction accuracy across 200 data subsamples, while adjacent asterisks represent significant clusters of classification above chance from 1000 permutations (Methods).

The prediction accuracy of novel v. repeat for subclip 1 was significantly above chance from 500-2500 ms after video onset (Fig. 1A), indicating the population of hippocampal neurons can differentiate new from old videos via a rate code. This period of classification matched the 2 s length of subclip 1, albeit with an offset from image presentation of ~500 ms, close to the 300-400 ms latencies typically shown in human hippocampal recordings to naturalistic stimuli(Quiroga, 2012).

For the analysis just described, we used all three trial types (Fig. 1A) to train and

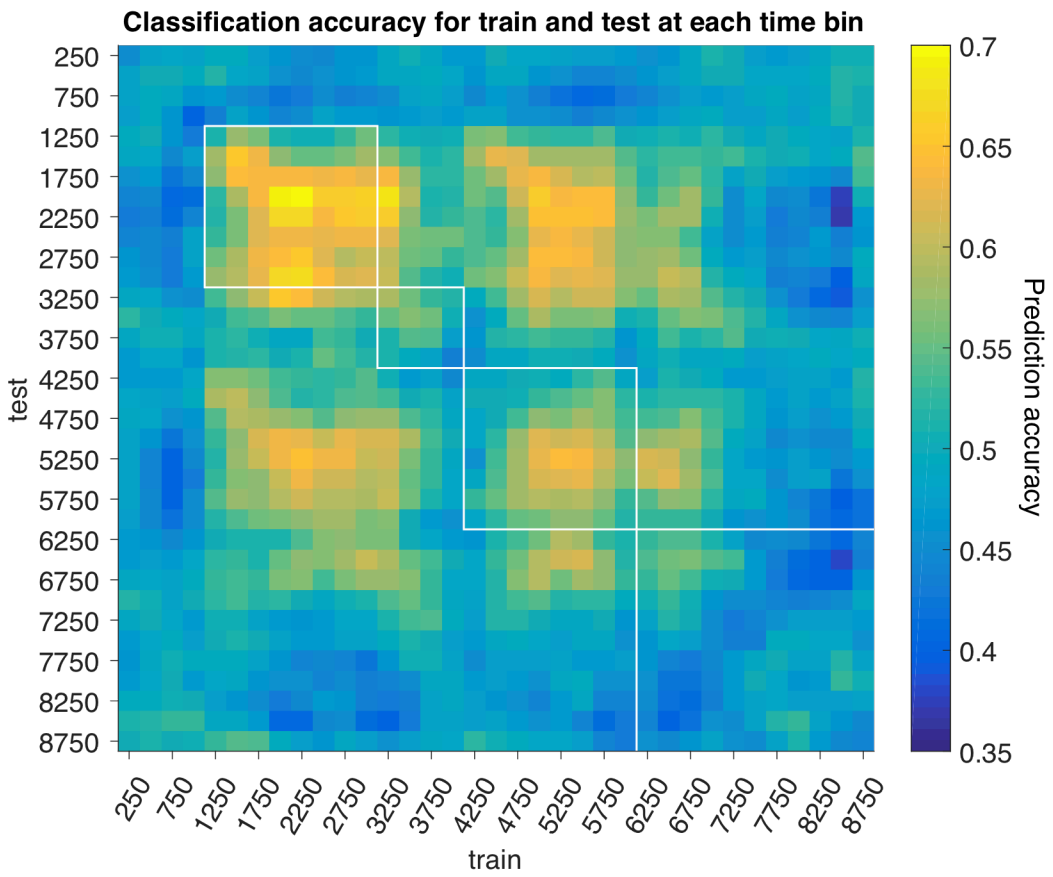


Figure 3. Training and testing classifier at different times reveals a stable neural code. Using the same classifier detailed in Fig. 2, prediction accuracy of novel v. repeat videos was assessed via 10-fold cross-validation at each point in the heatmap by training on one time bin and testing on the other. White boxes—in order left to right—indicate subclip 1, delay, subclip 2, and ITI periods. The diagonal of the heatmap is equivalent to the prediction accuracy in Fig. 2B.

test the classifier. And since we trained the classifier for subclip 1—where we used the predictor "repeat" for the first two trial types and "novel" only for the third trial type—we cannot simultaneously classify subclip 2 since in this case the predictor would be "repeat" only for the first trial type. Therefore, to use the same classifier for both subclip 1 and subclip 2, we removed trial type 2, thereby only classifying trials where the monkey was shown either repeat+repeat or novel+novel for subclips 1+2. Despite using only these 2/3rd of total trials, the prediction accuracy of novel v. repeat was significantly above chance for clusters in both the subclip 1 and subclip 2 periods (Fig. 2B). Classification across both periods using the same classifier suggests a similar rate code underlies repeat video recognition in each period.

Next, since the classification for the subclip 1 and subclip 2 periods were both novel v. repeat, we can essentially double our trial count if we pool both periods together and run the same novel v. repeat classifier. In this case, prediction accuracy approaches 80% at its peak, ~10% higher than the accuracy when looking at only subclip 1, and once again is significant from 500-2500 ms after video presentation

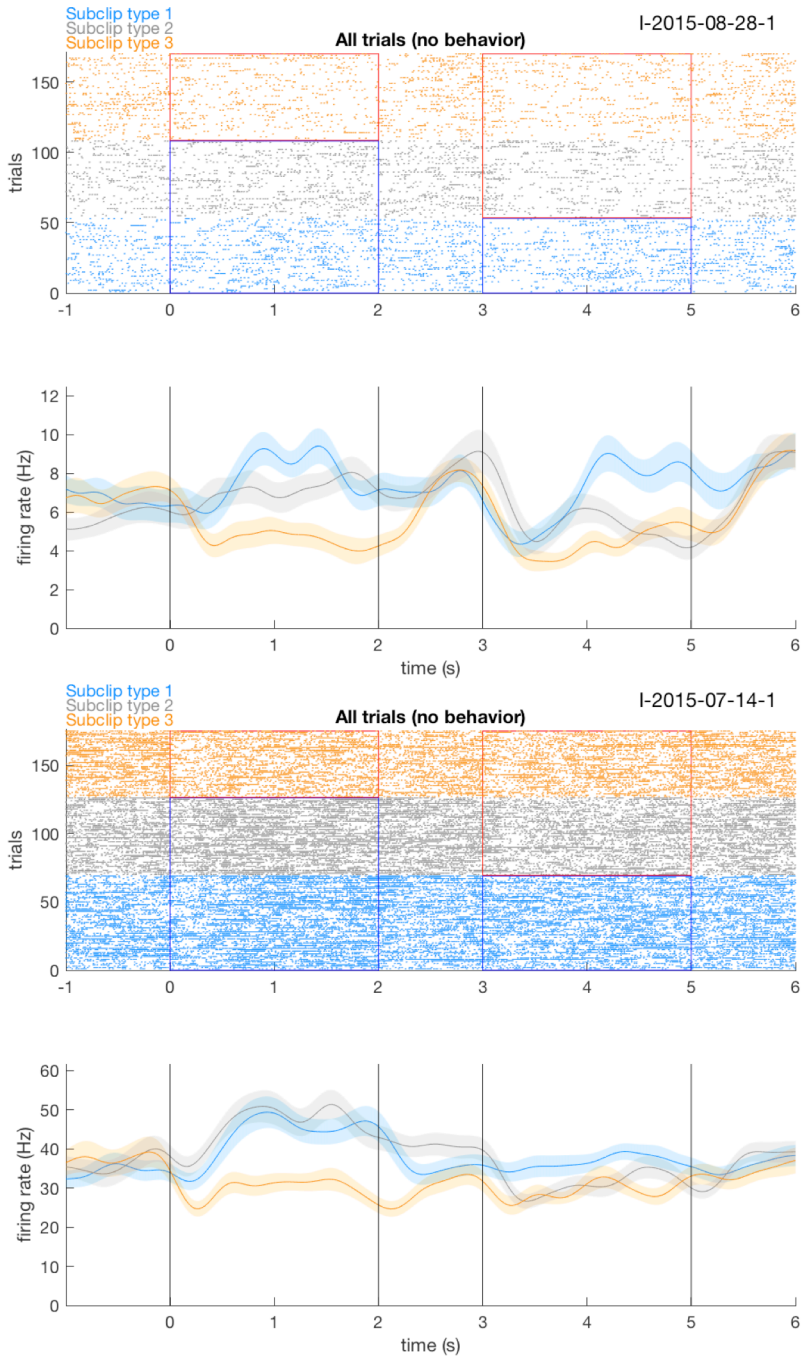


Figure 4. Example neurons. Rasters (top) and probability density functions (PDFs, bottom) for two different neurons. Trials are broken up by subclip type (1: novel/novel, 2: repeat/novel, and 3: repeat /repeat from top for subclip 1/2) and aligned to subclip 1 on. Vertical black lines are boundaries of subclip 1 (2 s), delay (1 s), and subclip 2 (2 s). In the rasters, areas outlined in red indicate presentation of novel videos, while areas outlined in blue indicate repeat videos. For example, subclip type 2 trials in gray are when monkeys are shown a repeat video for subclip 1 and after the delay a novel video for subclip 2. Dark lines in the PDF are average firing rate and outlines are SE.

(Fig. 2C). Stronger classification when pooling trials across the two subclips further suggests that neurons utilize a similar code for repeat video recognition across the two repeat v. novel conditions.

While these last two analyses imply that the population of hippocampal neurons classify novel and repeat videos via the same rate code both 1) within each 2 s subclip period and 2) between the separate subclip 1 and subclip 2 periods, since we train and test our classifier at each time bin, it's possible different sets of neurons or the same neurons with different signals over time are responsible for the clusters of classification we see in Fig. 2. To show that the same set of neurons use a stable rate code over time, we can train the classifier on one time bin and then test (cross-validate) the classification on a different time bin(Stokes et al., 2013). This analysis can

address 1) similar coding across each subclip, since we can train and test time bins within the same subclip period, and also 2) between the two subclips, since we can train in a subclip 1 time bin and test in subclip 2 time bin (and vice versa). Once again, as in the analysis for Fig. 2B, we only use the 2/3rd of trials where the monkey saw repeat+repeat or novel+novel for subclips 1+2. A heatmap of the classification accuracy when we train and test at each time bin shows evidence of a stable neural code both within and between subclips (Fig. 3). Prediction accuracy approaching 70% can be achieved 500 ms into the subclip 1 period (first white box) and shows approximately equivalent accuracy for all combinations of training and testing within the next 1500 ms. A similar phenomenon is present within the subclip 2 period, albeit with prediction accuracies only approaching 65% as in the subclip 2 period of Fig. 2B. Meanwhile, stable regions of prediction accuracy occur if we train on the subclip 1 period and test in subclip 2 or train on the subclip 2 period and test in subclip 1, indicating that the same set of neurons contribute to the novel v. repeat classification in both periods.

Next, we investigate the individual neurons that contribute most to these classifiers to understand the neural code underlying recognition of novel v. repeat videos. Two example neurons are shown in Fig. 4 (4 additional examples are shown in Fig. S2-S3), which represent the neurons with the 2nd and 12th most positive weights in the classifier for the subclip 1 period and the 5th and 12th most positive weights in the classifier for the subclip 2 period (Fig. 2). In these cases, positive weights are indicative of classification via higher firing rates to repeat than novel videos. Therefore, negative weights would indicate higher firing rates to novel than repeat videos, while weights of 0 mean no difference in firing rates between novel and repeat videos. The distribution of weights across the population for the classifiers for the two subclip periods—after setting the first neuron's weight to 1 and normalizing the remaining neurons based on this value—are shown in Fig. 5A-B (red).

The majority of neurons that contribute to the classifier have positive weights, with firing rates to repeat videos greater than novel videos as in the examples of Fig. 4. In fact, when we train a classifier across the entire subclip 1 period (from 500-2500 ms after video presentation), of the 119 neurons in this population, 23 had weights greater than 0.3, while only 5 had weights less than -0.3. This distribution is in stark contrast to our previous hippocampal recordings in monkeys during a serial image task, where animals viewed sequences of images that contained novel, repeat, and lure (similar) images. Plotting normalized weights from the classifier in that task, in which positive also indicates repeat > novel firing rates and negative indicates novel > repeat firing rates, we see a much more even distribution of positive and negative weights (Fig. 5A, black). In that case, of the 114 neurons in the population, 15 had weights greater than 0.3, while 19 had weights less than 0.3. The difference between these two distributions is significant ($p=0.0013$, Wilcoxon rank-sum test). We repeated this analysis for the subclip 2 period from 500-2500 ms after video presentation, and show a similar result (Fig. 5B, $p=1.1E-4$, Wilcoxon rank-sum test).

In sum, while our previous work(Sakon and Suzuki, 2019) (Fig. 5A, black) and others(Jutras and Buffalo, 2010; Rutishauser et al., 2006; Rutishauser et al., 2008) have shown a similar mixture of neurons that code for novel v. repeat recognition via either novel>repeat or repeat>novel firing rates, in our video task hippocampal neurons predominantly show this latter code (repeat>novel firing rates) as in the examples in Fig. 4.

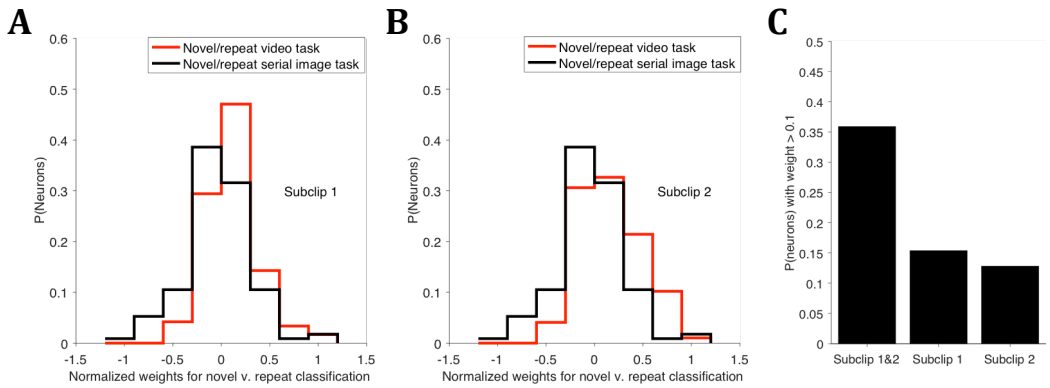


Figure 5. Distribution of weights contributing to population classifier. **A)** The red line shows the neurons for the classifier trained and tested on subclip 1 as in Fig. 2A, but for a single time bin across the subclip 1 period. The black line is for the same classifier trained and tested on hippocampal neurons recorded during a task where monkeys viewed serially presented novel and repeat images(Sakon and Suzuki, 2019). The weights for each neuron were ranked and normalized such that the most positive weight in each distribution was set to 1. The weights shown in red here are plotted by anatomical location in Fig. S1, which shows the neurons that contribute strongly to the classifier are found throughout hippocampus. **B)** Similar to A, but comparing the same classifier trained and tested on the subclip 2 period with the serial image task. **C)** The proportion of neurons with weights for subclip 1 and subclip 2 that were >0.1 (i.e. 10% of the best classifying neuron), using the classifiers for subclip 1 and subclip 2 shown in A and B. Subclip 1&2 indicates the weight for those neurons was >0.1 for both periods, while neurons with weights >0.1 for only one of the periods is shown in the next two bars.

Finally, we return to the question of whether these hippocampal neurons show a general recognition memory response across both subclip periods of the task. In other words, are the same neurons responsible for repeat v. novel video classification in both subclips, or do some neurons differentiate these two "contexts" by classifying within only one subclip? Earlier, we found that of the 36 neurons with significantly higher firing rates to repeat than novel videos in subclip 1 and the 34 neurons with significantly higher firing rates to repeat than novel videos in subclip 2, 22 of these neurons are the same. Now, with the use of the classifier, instead of relying on multiple significance tests to group neurons, we can assess the ubiquity of each neuron's recognition memory response by comparing its weights to the separate classifiers trained on subclip 1 and subclip 2 shown in Fig. 5A-B. Indeed, of the neurons with weights >0.1 in at least one of the periods, 56% had weights >0.1 for both periods (Fig 5C), indicating higher firing rates to repeat than novel videos. Meanwhile, only 24 and 20% of neurons showed weights >0.1 for the subclip 1 and subclip 2 periods, respectively. Therefore, the majority of recognition memory

responses—shown via higher firing rates to repeat than novel videos—are general across contexts.

Recall is not reflected by hippocampal neurons during the delay. A second key expectation from our task design is that recall of clip b (Figs. 1A and 6A) will be represented in the neural activity during the delay period after subclip 1 (note that, for ease of explaining our comparisons in the following section, we refer to 0-2 s, 2-4 s, and 4-6 s of each clip as clips a, b, and c (Fig. 1A); e.g., the first part of the first clip is clip 1a) Once the monkeys view the 6 s clip three times during the encoding period, they are shown the first 2 s of that clip on the 2/3rd of trials when subclip 1 is a repeat (Fig. 1A). We anticipate this 4th presentation of the first 2 s of the clip during the subclip 1 period will elicit recall of the video by the monkey, thereby cuing the hippocampus to recall the next segment of the video (i.e., clip b) during the video-free delay.

We explored two analyses that might show evidence of recall during the delay. For the first analysis, we took each of the same 110 neurons used in our classifier, and found a vector of firing rates during the delay period for each trial after a repeat was shown in subclip 1. Next, since we expect this 4th repeat of the same video to elicit recall of the remainder of the clip, we found a vector of firing rates for those same trials during clip 1b, and then correlated these vectors of firing rates for each neuron (Fig. 6A). For comparison, we correlated the clip 1b vector with the delay period on trials when subclip 1 was a novel video, in which case we do not expect recall of the remainder of the clip to be elicited. Our expectation is for higher correlations of clip 1b v. delay when a repeat video instead of a novel video is shown during subclip 1.

Contrary to our expectation for recall, the correlation between clip 1b and delay after presentation of repeat videos was no different than between clip 1b and delay after novel videos during subclip 1 (0.15 ± 0.02 v. 0.17 ± 0.02, Fig. 6; full distribution shown in Fig. S4). We found the same trend when we correlated clip 3b to the delay period as well, with lower correlations after repeat videos compared to after novel videos (0.14 ± 0.02 v. 0.16 ± 0.02, Fig. 6). These four vectors of correlations are not significantly different than if they were drawn from the same distribution ($p=0.83$, multilevel ANOVA). In fact, when we correlate clip 1a with subclip 2—two periods that never show the same video and therefore should not correlate—the value is similar to these other comparisons (0.13 ± 0.02, negative control in Fig. 6), indicating the delay period shows no evidence of recall beyond basal levels of correlated firing.

For our second analysis of recall, we considered if a repeat video during subclip 1 elicits something more like a prospective signal(Lisman et al., 2017) for subclip 2. In other words, after the repeat cue in subclip 1, hippocampal neurons might show evidence of recalling subclip 2 during the delay. To test this, for each neuron we measured the correlation in firing rates across trials between the delay period after subclip 1 was a repeat v. subclip 2. If a prospective recall signal for subclip 2 exists,

the correlation should be stronger when subclip 2 was then a repeat than when subclip 2 was a novel video. Instead, these two correlations were not significantly different across neurons (0.24 ± 0.02 v. 0.19 ± 0.02, p=0.10).

The lack of recall signal in this analysis could be due to a number of factors. First, trial-by-trial correlations of firing rates might be a poor metric for comparing the state of the hippocampal population. To test this idea, we can correlate the firing rates of the same clip across different periods (clip 1a v. clip 3a; positive control between clips in Fig. 6) or the firing rates at different times within the same clip (clip 1a v. clip 1c; positive control within clips in Fig. 6). The correlations within hippocampal neurons for these two positive controls were higher than the other four distributions, and were not expected by chance when compared to the other distributions (p=0.017, multilevel ANOVA). Therefore, the hippocampal population does show trial-by-trial evidence of returning to a similar state, but only during the same encoding video (clip 1a v. clip 3a) or within the same temporal context (clip 1a v. clip 1c).

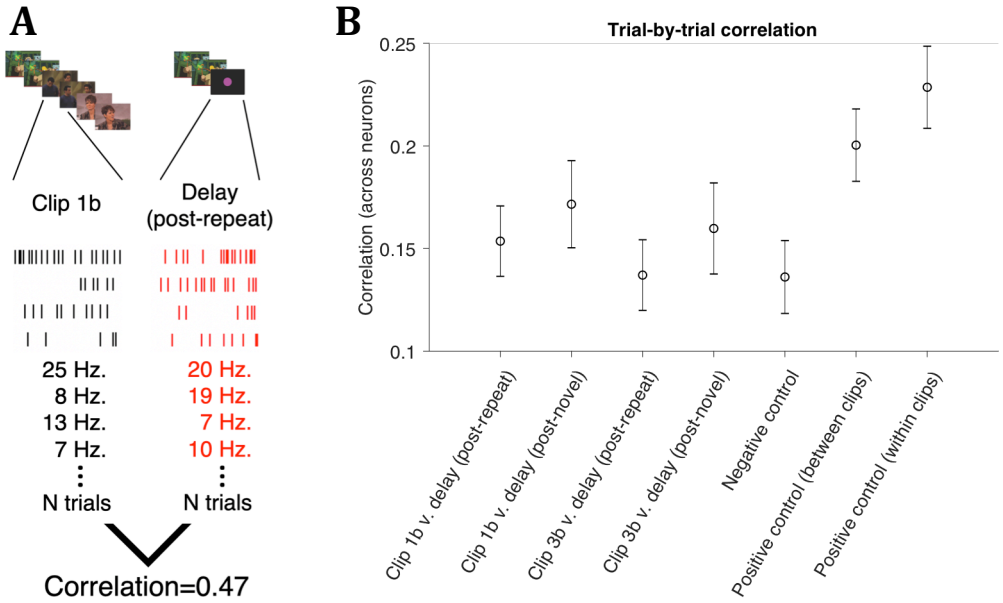


Figure 6. The hippocampal population only shows correlated firing during the clip periods. **A)** Example correlation (Spearman) analysis for a single neuron. To test cued recall, vectors of firing rates during key periods for each neuron were correlated to find trial-by-trial evidence of the hippocampal population returning to a similar state. The expectation is that if the monkey is cued to recall the clip video, delay period firing rates after a repeat video is shown in subclip 1—but not after a novel video is shown in subclip 1—should correlate with the firing rates to the middle video of the clip. **B)** Average correlations across neurons. The negative control is the correlation between unrelated videos, while the positive controls are the correlations between repeated clips or within the same clip. Error bars are SE. Full distributions of neurons are shown in Fig. S4.

Finally, since we see no evidence of recall, we asked if the recognition signal shown in Figs 2-3 might carry over into the delay period instead. We used the classifier from Fig. 2, since it was trained on identifying novel v. repeat videos from neural

responses, but focus on neural responses aligned to delay onset (Fig. 2D). Instead of classification continuing into the delay period, the prediction accuracy for whether a repeat or novel video was just shown during the subclip 1 period quickly dissipates, with the end of the cluster of significant classification 300 ms after delay onset (Fig. 2D; note the use of smaller integration windows for better resolution). The same dissipation can be seen at the level of individual neuron firing rates as well, with the differences in firing rates between novel and repeat videos quickly disappearing upon delay onset in all six of the example neurons that contribute strongly to the subclip 1 classifier (Figs. 4 and S2). Considering the time lag in the recognition memory signal after stimulus onset, where significant classification only begins 400 ms after subclip 1 videos are presented (Fig. 2D), the classifier does not show evidence that the hippocampal population carries a signal for recall of repeat videos into the delay period. Overall, these results suggest that hippocampal neurons did not show recall or hold recognition memory of the previous video into the delay, likely due to inadequate task demands for the monkeys to hold memory into this period as explained below.

Discussion

We adopted a task used in humans(Gelbard-Sagiv et al., 2008) for non-human primates to separably investigate recognition and recall memory and electrophysiologically recorded from the hippocampus to assess how single neurons in this structure reflect each form of memory. Animals were engaged in the task, as their looking times indicated significantly greater interest in novel than repeat videos. Crucially, a large proportion of hippocampal neurons discriminated the presentation of novel or repeat videos during the two key retrieval (subclip) periods, with the dominant signal for repetition surprisingly being an enhancement in firing rates. Most neurons that reflect this repetition enhancement signal in the first retrieval condition also show this same signal in the second retrieval condition, indicating a general code for recognition memory in hippocampus. Finally, we find no evidence of hippocampal neurons reflecting cued recall during the delay period, likely due to insufficient task demands in our free-viewing paradigm. These results provide evidence for hippocampal involvement in implicit recognition memory of episodic stimuli and suggest repetition enhancement as a signature of this memory at the level of single hippocampal units.

The majority of recognition memory studies recording single hippocampal neurons in monkeys have required explicit identification of repeated stimuli—typically via a touch(Brown et al., 1987; Riches et al., 1991; Xiang and Brown, 1998) or lick (Rolls et al., 1993). These studies found few (<5%(Rolls et al., 1993)) or no(Brown and Xiang, 1998) hippocampal neurons significantly responded to repeated stimuli. However, recent studies(Jutras and Buffalo, 2010; Sakon and Suzuki, 2019) using implicit paradigms—where monkeys were serially presented images and not required to indicate repeats—have shown large proportions (~50%) of hippocampal neurons significantly change in firing rate to repeat images (as summarized in Table 1). Hippocampal recordings had not been performed while

monkeys watch relatively natural video stimuli, making it unknown if the recognition memory signals during these tasks were limited to still images or serially presented stimuli (where monkeys controlled how long the images appeared onscreen in both studies). Here, we find a significant proportion of hippocampal neurons that show a recognition memory response via changes in firing rates to repeated stimuli, although the majority of video-responsive neurons showed repetition enhancement to repeat videos (~70% of neurons to either subclip 1 or subclip 2, Table 1), whereas the two serial image studies showed predominantly repetition suppression (~60% of neurons in both studies, Table 1). Since this is the first study using video stimuli, it is unknown if these different signatures of recognition memory are driven by the implicit, fixed-time nature of the task or the use of more natural stimuli.

While monkey electrophysiology studies using explicit recognition memory tasks have shown little evidence of hippocampal involvement, a number of human intracranial electroencephalography (iEEG) studies using such tasks have shown evidence of recognition memory via changes in hippocampal firing rates to repeated images. One study showed repetition enhancement in hippocampus to correctly-identified, repeated words(Merkow et al., 2015), although this study only used high-frequency activity as a proxy for firing rates and therefore did not have single neuron resolution to contrast repetition enhancement and repetition suppression as in our work. Another pair of studies showed evidence of recognition memory in humans via both repetition suppression and repetition enhancement neurons(Rutishauser et al., 2006; Rutishauser et al., 2008), although the task design in these studies involved an object-in-place(Wirth et al., 2009) learning paradigm, which is likely to engage the hippocampus more via its clear role in associative learning(Wirth et al., 2003) than recognition memory per se. Another study had humans identify repeated images in serially presented blocks of faces or scenes, and found a large proportion of single hippocampal units that responded to previously seen stimuli(Viskontas et al., 2006). In this study, 69% of these hippocampal neurons showed repetition suppression, which is in line with the serial image monkey tasks described in the last paragraph. Considering the clear divide in monkeys for explicit recognition tasks showing no hippocampal responses and implicit recognition tasks (as shown here) showing strong changes in firing rate, it is difficult to reconcile this study with monkey work, although differences in how monkeys and humans utilize familiarity v. recall (e.g. recall-to-reject) to solve explicit tasks might underlie the discrepancy.

In addition, while to the best of our knowledge video stimuli have not been used in electrophysiological studies of monkey memory, videos have been used in a human electrophysiological (iEEG), single-unit study of memory(Gelbard-Sagiv et al., 2008). However, the paradigm in this work selected for neurons responsive to long-term memory "concepts", where videos of people (e.g. Tom Cruise) or content (e.g. The Simpsons) meaningful to the subjects are screened and single neurons selectively rise in firing rate while viewing or later recalling those specific videos. In our results, the hippocampal neurons respond stronger to repeated videos agnostic to

the content of the individual videos, as the change in firing rate for these neurons is clear after averaging across trials (Fig. 4), even though unique videos with variable content were used on every trial throughout all recording sessions. The two novel v. repeat task conditions during the retrieval period further confirm the generality of this recognition memory signal, as the classifier trained on hippocampal population responses in one subclip period can decode which trials shown to the monkeys were repeats in the other period (Fig. 3), even though the videos shown in subclip 1 and subclip 2 were never the same since the first and last two seconds of the clip were from different sources. In fact, adding to the generality of the recognition signal, the neurons responsible for classification via repetition enhancement in one subclip period tended to be the same ones responsible for classification in the other (Fig. 5C).

We attempted to emulate one other aspect of this human iEEG video study, where patients would freely recall videos they had just seen, and the same hippocampal neurons that previously increased in firing rate to those videos would reactivate when the patients audibly recollected the videos(Gelbard-Sagiv et al., 2008). Since we could not ask the monkeys to freely recall videos they had just seen, we tried repeating the clips 3x to promote strong encoding, and then cueing the monkeys to recall the clip by repeating the first part of the clip video (0-2 s) during subclip 1 followed by a delay period showing only a static dot. Then, by contrasting neural responses during this delay to when a novel video was shown in subclip 1 instead of a repeat (the novel video would not cue recall since it is always a true first presentation), we hoped to show neural evidence of hippocampal reactivation as monkeys recalled the remainder of the clip. Instead, we found no evidence of hippocampal reactivation, as the classification of repeat v. novel videos quickly dissipated once the delay period started (Fig. 2D), and the hippocampal population failed to reactivate into a similar firing rate state during the delay as in the clip (Fig. 6). Since the hippocampal population did show similarity in firing rates within the same clip and between the identical clips (positive controls, Fig. 6B), we suspect our task failed to elicit recall in the monkeys. This might not be surprising, since the monkeys had no reward incentive to keep the clip periods in mind during the delay period. In addition, owing to the attention span of the monkeys, we made the delay period only 1 s between the subclip periods, while the structure of the clips (three contiguous, 2 s videos) suggests a 2 s delay would have been a better match. This mismatch also could have disrupted recall. Overall, while hippocampal neurons show clear evidence of recognizing repeated videos while they are presented (Fig. 2-4), to test cued recall in monkeys, hippocampal neurons would have to be screened more like in human concept cell studies and then probed with their selective stimuli. Since we used unique videos for every trial, we do not have enough trials with the same stimuli to find such concept cells in our population. One study in monkeys has shown evidence of hippocampal neurons tuned to conceptualizations of a virtual environment(Baraduc et al., 2019), suggesting that cued recall of concepts more like those shown in human concept cell studies is possible in monkeys.

The involvement of the hippocampus in recognition memory has been hotly

debated, with a recent proposal suggesting the hippocampus is recruited in recognition memory when stimuli are repeatedly used across experiments(Bird, 2017). We make the case above that, since we used trial-unique stimuli in both our previous work(Sakon and Suzuki, 2019) and this work, the hippocampus is also involved in recognition memory of novel stimuli during tasks without an explicit response. However, unlike previous work where neural responses in the hippocampus showed an approximately even mixture of suppression and enhancement to repeated stimuli—much like has been shown during DMS tasks in perirhinal(Miller et al., 1991) and entorhinal cortex(Suzuki et al., 1997)—neurons recorded during repeated videos in our task show a majority of repetition enhancement responses (Table 1-2, Fig. 5). We suggest that the use of video stimuli in our task might be responsible for this previously unseen pattern of responses. Videos contain series of object, faces, and scenes within a context, and using such life-like stimuli as typically used in human studies of hippocampus(Baldassano et al., 2017; Ben-Yakov et al., 2013; Chadwick et al., 2010; Gelbard-Sagiv et al., 2008) are likely to yield neural correlates indicative of the naturalistic experience of everyday life(Krakauer et al., 2017). In addition, videos introduce an element of time, and hippocampal neurons have shown strong evidence of roles in elapsed time(Kraus et al., 2013), temporal order(Naya and Suzuki, 2011), and associative memory tasks with a timing condition(Sakon, 2014).

In 1986, Marvin Minsky wrote "we're least aware of what our minds do best," a concept since termed Moravec's paradox(Minsky, 1986). The crux of the paradox is that many things that come to us effortlessly—like being able to recognize moment-to-moment changes of objects and environments—are largely done without conscious thought. Meanwhile, the brittleness of memory search—a tradeoff between storing new information of objects or environments while still being able to efficiently generalize and index those memories at a moment's notice—is known to be a difficult problem for computers(Lenat et al., 1990; Marcus, 2018). Here, in a task where monkeys free-viewed naturalistic videos without having to make explicit responses for reward, thereby avoiding a requirement of the vast majority of memory tasks to date, we found strong evidence of hippocampal involvement in recognition memory, and in a repetition enhancement pattern previously unseen in single unit memory studies. Therefore, we suggest the hippocampus is recruited in the counterintuitively effortless yet computationally difficult task of implicit recognition memory. This finding uncovers a largely unexplored direction in determining how the brain handles both implicit and explicit recognition memory, and understanding this process might contribute to computational algorithms to improve generalization problems in memory search(Marcus, 2018).

Methods

Subjects: Two female rhesus macaque (I = 6.1 kg, age 7; W = 6.5 kg, age 11) monkeys were surgically implanted with titanium headposts using aseptic techniques under general anesthesia and trained to be comfortable visually interacting with a computer monitor for juice reward while fixed in a primate chair for multiple hour recording sessions. All procedures and treatments were done in accordance with NIH guidelines and were approved by the New York University Animal Welfare Committee.

Task: The task involved presentation of a series of videos with fixation "games" that rewarded the monkey before each video was shown. The game was to center fixate on a purple dot (radius = 1°, fixation maintenance required within 1.3° of dot center for 500 ms) to earn drops of 50% juice reward. Reward was never given during video presentation in order to separate reward from task-relevant stimuli. Monkeys were only required to fixate their eyes during these fixation games and during pre-video periods, which occurred before the start of videos to ensure the monkeys' attention was drawn to the center of each video before it was presented. Pre-video dot properties were identical to fixation game dots except the color was changed to pink. Once pre-video fixation was achieved a video was shown (either clips 1-3 or subclip 1), and monkeys were permitted to freely view them, thereby allowing us to use natural viewing as a gauge of their interest in a given video (preferential looking).

The encoding phase of each trial began with a fixation game, where monkeys earned 0.35mL of 50% juice for achieving fixation. Next, a pre-video dot appeared, and once the monkey achieved fixation on this dot for 500 ms, clip 1 was shown. Then, after a second fixation game and pre-video fixation, the same 6 s video was repeated as clip 2. Finally, after a third fixation game and pre-video fixation, the same 6 s video was repeated as clip 3. A blank screen was shown for 450 ms after clips 1-2 before the dot reappeared for the next fixation game. A blank screen was shown for 300 ms after clip 3 before entering the retrieval period.

Clips were comprised of 3, 2 s videos shown contiguously as a single, 6 s episode (19.1° x 14.3° [640 x 480 pixels]). We refer to the 3 contiguous videos that comprise each clip as a, b, and c. Clips 1, 2, and 3 were identical videos meant to improve encoding through repetition. The videos were drawn from a new set of >40, 1-3 min source videos each daily session and presented without sound. The source videos were taken from Internet videos (YouTube) of scenes involving animals, people, animations, and other such moving subject matter. Music videos with a changing narrative (not focused on the same scene) and trailers to movies were typically used due to their short length and fast-moving, non-repetitive content. A custom Presentation (Neurobehavioral Systems) program was written that would randomly select 2 s snippets from that session's source videos, only take snippets at least a 5 s buffer from any previously seen snippet, and ensure each snippet would never be used again. While there was a chance of similar clips across trials since we used the same source videos each session, the short scenes of the music videos/trailers as well as the use of the buffer minimized overlapping content. In addition, since each clip was always comprised of 2 s snippets from 3

different source videos, the clip itself was almost guaranteed to be a unique construction of content sources since even with only 40 source videos there are 40^3=64,000 order arrangements, while monkeys averaged only 125.9±2.5 trials per session.

The retrieval period of each trial began with a fixation game and a pre-video period before subclip 1 was shown. Immediately after subclip 1 was shown, a larger (1.6°), purple dot was shown in the center of the screen for 1 s. The monkey was not required to fixate to this dot, which served merely as an indicator of the delay period between subclips. Finally, subclip 2 was shown followed by a 3 s inter-trial interval.

Subclips were comprised of 2 s videos surrounding the 1 s delay period. Three different trial types with a different combination of subclips were shown at equal proportions: i) subclip 1 a repeat of the first 2 s of the clip (i.e., the first of the three snippets in the clip) and subclip 2 a repeat of the last 2 s of the clip, ii) subclip 1 a repeat of the first 2 s of the clip and subclip 2 a novel (unseen) video snippet from the source videos, and iii) subclip 1 a novel video and subclip 2 a novel video. The delay period was only set at 1 s—instead of 2 s to better match the structure of the clips—owing to pilot behavioral data indicating delays longer than 1 s made it more difficult for monkeys to maintain attention.

Behavior: We used looking times at the videos as a proxy for memory of the repeated videos (preferential looking). Monkey eye scan paths were recorded using an infrared camera (IScan, Inc.) as the animals viewed a 19-inch LCD monitor (Dell) positioned 0.55 m from the left eye to the center of the screen. Monkeys were considered to be looking at the video when their eyes were within the bounds of the 640x480 pixel video plus a 20 pixel buffer surrounding the perimeter of the video. We then measured the proportion of the time that the monkey looked within the bounds of the clip or subclip videos to find looking times. Receiver operating characteristic (ROC) curves were then created for each session with the threshold parameter set to the % of looking time between a novel and repeat video from the range 0-125% with 2.5% step size. When the monkey looked longer at novel subclips than clip*threshold, we considered this a true positive. Therefore, when monkeys looked longer at repeat subclips than clip*threshold, this was a false positive. We found the threshold for each session that maximized the true positive rate while minimizing the false positive rate, and calculated d' values for each session by subtracting the z-score for the false positive rate from the z-score of the true positive rate.

To give a concrete example, if the threshold is set to 80%, the monkey looked within the bounds of a clip for 1.6 s, and the monkey looked within the bounds of a novel subclip for 1.4 s, this subclip would be a true positive since 1.6*0.8<1.4 s. However, if the threshold is set to 100%, holding the numbers the same, this subclip would be a true negative, since 1.6*1.0>1.4 s. Trials were only considered for analysis if the clip period was looked at for at least 0.5 s of the 2 s segment.

We performed two comparisons to analyze behavior: clip 1a (0-2 s) v. subclip 1, and clip 1c (4-6 s) v. subclip 2, since these were the two segments of clip 1 that videos could be

repeated in the retrieval period. Average thresholds for the 110 sessions was 88.0±1.6% for clip 1a (0-2 s) v. subclip 1 and 86.5±1.6% for clip 1c (4-6 s) v. subclip 2. Results were similar when we used clip 3 instead of clip 1 in these comparisons. Statistics for comparisons of average looking times (Fig. 1B) for each monkey were Bonferroni-corrected for family-wise error rate (FWE) by adjusting the p-values for the 6 total comparisons (6*p-value).

Electrophysiology: Once monkeys were trained on the task, MRI-designed, custom-fit circular HDPE plastic recording chambers (Rogue Research) were cemented to their skulls with acrylic perpendicular to the horizontal plane in stereotaxic coordinates and dorsal to their right hippocampi (chamber outlines are shown in Fig. S1). Single-unit recordings were performed by lowering glass-insulated electrodes (~1-MΩ impedance; Alpha Omega) or glass-insulated tetrodes (~0.5-MΩ impedance for all four shafts; Thomas Recording) via a hydraulic (Kopf Instruments) microdrive through 23-gauge metal guide tubes that reached from the chamber grid to ~10–15 mm dorsal to hippocampus. Guide tubes were placed at new locations at the beginning of each week and removed at the end of the week after multiple sessions of recording. Neural signals were acquired, filtered, amplified 1,000 times, and digitized via the MAP Data Acquisition System (Plexon) at 40,000-kHz resolution. Single unit waveforms were sorted manually using a combination of offline sorter (Plexon) and superparamagnetic clustering [Wave_Clus(Quiroga et al., 2004)].

Single unit analysis: We recorded a total of 209 neurons across the 110 sessions detailed in our behavioral analysis. All of these neurons were used in our population sparseness analysis, which was performed identically to previous work(Sakon and Suzuki, 2019). Briefly, population sparseness is a measure of the proportion of neurons in a population that respond to a given stimulus. Ideally, this measure is done for single stimuli repeated multiple times across trials, but since by design we do not repeat stimuli across trials, we measured the sparseness across videos shown during subclip 1 for all trials (therefore, instead of pooling trials for the same stimulus, we pooled trials for the same task period). The population sparseness value of $a^{\mathrm{P}} = 0.30$ was similar to values for monkey hippocampal neurons firing to repeated stimuli ($a^{\mathrm{P}} = 0.33$) in previous work (Rolls and Treves, 2011).

To perform the single unit analyses in Tables 1-2, we kept only those neurons recorded for at least 20 trials and with a minimum average firing rate of at least 0.5 Hz across all trials. The number of neurons used in each analysis are listed in the table. Spikes in the window of 0.1-2.1 s were measured separately for each condition (novel and repeat). Neurons with significant increases or decreases in firing rates (Table 1) were found by performing a Wilcoxon rank-sum test ($p<0.05$) between trials from a given condition and firing rates pooled across all fixation games (there are 4 such fixation periods per trial—before each clip and subclip 1). Neurons with significant differences between conditions (Table 2) were found using the same test. Chi-square tests were used in both tables to assess the likelihood of one group of neurons being larger than another (e.g. increase v. decrease in firing rate to repeats), with p-values being Bonferroni-corrected to account for family-wise errors (FWR) by the number of tests within each table (i.e., 4 tests for statistics in Table 1, and 2 tests for statistics in table 2).

Population classifier analysis: We performed a 10-fold, cross-validated, L2-regularized (ridge) logistic regression for the neural firing rates in our population with fixation period (baseline) firing rates >0.5 Hz. This setup means the classifier was trained on 90% of trials and tested on the remaining 10%, and this process was repeated 10x in a balanced design such that every trial was part of the test set once. Predictors were the identity of the subclip 1 video (novel or repeat) with the response variable the firing rates for each neuron. Separate classifiers were trained on each time bin, with 500 ms spike integration windows and 250 ms step sizes for Fig. 2A-C and 200 ms integration windows and 100 ms step sizes for Fig. 2D. We used regularization since with small numbers of trials on each run perfect classification on any given training or test set would weight that neuron to infinity. Ridge regression was used since our goal was to find all neurons that contributed to classification and not remove redundant signals. By keeping weights for each neuron, we could also gain an approximate rank order of each neuron's contribution to the classifier (Fig. 5). The regularization parameter (λ) was set to 0.15, although a range of numbers (from 0.05 to 0.3) yielded similar results.

We attempted to use a similar number of neurons in each classifier we trained, with the number of neurons differing depending on the minimum number of novel and repeat trials we required for each analysis. For the classifier of subclip 1 identity in Fig. 2A (and also 2D), where we used all trials, we required at least 40 novel and 40 repeat trials for each neuron, which gave a population of 96 neurons. For the classifier of subclip 1 and subclip 2 identity in Fig. 2B & Fig. 3, where we eliminated the (1/3rd of) trials which were repeat then novel for subclips 1 and 2, we required at least 35 novel and 35 repeat trials for each neuron, which gave a population of 86 neurons. Finally, when we combined trials across subclips 1 & 2, we required 100 novel and 100 repeat trials, which gave a population of 106 neurons.

The prediction accuracy for each classifier is an average over many permutations sampled from different trials in the same group of neurons. Since the response variable for the classifier trained at each time bin is a neuron X trial matrix (e.g., for Fig. 2A, 96 x 80), and many of those neurons will have more than the minimum number of trials we set for that classifier, we randomly subsampled trials for those neurons with more than the minimum. We ran the classifiers in Fig. 2A, 2B, and 2D for 200 permutations, averaged across those permutations to find the prediction accuracies at each time bin, and used the 5th and 95th percentile classification runs for the error bars shown in Fig. 2. For computational reasons due to the larger number of trials, we only used 80 permutations for the classifier in Fig. 2C. Notably, the prediction accuracies were remarkably consistent across multiple runs for all four classifiers.

To find periods of significant classification, we used a cluster-based metric that corrects for multiple comparisons across time bins (Maris and Oostenveld, 2007). For each of the 200 (or 80) subsampled permutations described in the last paragraph, we ran 5 additional permutations of the classifier with the labels (novel and repeat) shuffled, thereby creating a surrogate distribution of 1000 (400) total permutations. A maximum cluster was found for each of these surrogate permutations, where a cluster is the sum

of the prediction accuracies for a stretch of contiguous time bins that cross a threshold. The threshold can be set to any value without increasing the false alarm rate, but a threshold set too high or low will decrease the sensitivity(Maris and Oostenveld, 2007). We used 0.625 as the threshold for the classifiers in Fig. 2A and 2C, and 0.6 for the classifiers in Fig. 2B and 2D. Next, we took the real prediction accuracies, from averaging across the 200 (80) permutations, and found all contiguous clusters above threshold as we did for the surrogate data (in this case, not just taking the maximum cluster). Finally, we found which of these cluster values in the real data were greater than the 95th percentile cluster value in the surrogate data, and any such clusters above this value were considered significant. Significant clusters of classification in Fig. 2 are denoted with asterisks, with all asterisks in neighboring time bins indicating a single, contiguous cluster.

For the heatmap of prediction accuracy in Fig. 3, we trained and tested at each combination of time bins from 1000 ms before subclip 1 on until 3000 ms after subclip 2 off. The classifier at each point was identical to the classifier in Fig. 2B, with the diagonal of the heatmap being equivalent to the prediction accuracy in 2B. The prediction accuracies shown are the averages of 100 permutations with different subsamples of trials as described for Fig. 2B.

To find the neurons that most contributed to classification we see in the subclip 1 and subclip 2 periods (Fig. 4 and Fig. S2), we ran separate classifiers on a single time bin from 500-2500 ms after subclip 1 on and from 500-2500 ms after subclip 2 on. These weights provide the ranks of which neurons had the most positive weights that contributed to the classifier. The distribution of weights for the classifier for the two subclip periods also fill the distributions used in Fig. 5.

Cued recall analysis: We investigated neural responses during the delay between subclips with two analyses. First, we used a classifier in Fig. 2D similar to the classifier in Fig. 2B, but including all trials (the classifier in 2B removes the repeat/novel trial type for subclip 1/subclip 2). The predictor for this classifier was if the video in subclip 1 was a repeat or novel video, since we expect recall to only occur during the delay period after the monkey is cued with a repeat video—and not a novel video—in subclip 1.

Second, we looked at the trial-by-trial correlation (Spearman) of firing rates for each neuron when comparing the delay period with the portion of the clip we except to be recalled. That is, if a repeat video is shown during subclip 1 and this (4th presentation of 0-2 s of that video after the clip was shown 3x) successfully cues recall in the monkey, we expect neurons during the delay period to recapitulate firing rates on those trials. However, if a novel video is shown during subclip 1, that should not elicit recall, and we should not expect correlated firing between the clip and delay periods on those trials. For the delay period in these comparisons, we found the firing rate in the window from 0.1-1.1 s after subclip 1 off. For the clip period, we found the firing rate in the window from 2.1-4.1 s after clip on (since clip b should be what is being recalled after the monkey is cued with a repeat of the first 2 s of the clip, as diagrammed in Fig. 6A).

Statistical comparisons were all multi-level ANOVA, where we correlated out differences between monkeys and the interaction factor to isolate only the differences in variance between the correlations.

We measured 7 trial-by-trial correlations of firing rates this way and averaged them across neurons for Fig. 6B. For the first two comparisons, we correlated the middle 2 s of clip 1 with the delay period separately after a repeat video (post-repeat, when we expect recall) and after a novel video (post-novel, when we don't expect recall). We did the same comparisons for clip 3 v. delay as well.

Next, we ran 3 controls to get an idea of the correlation values we should expect across our hippocampal population. First, a negative control where we compared firing rates to videos in unrelated periods. We found the firing rates from 0.1-2.1 s after presentation of clip 1 to firing rates during subclip 2, where a video is presented that is never a repeat of the first 2 s of the clip (since subclip 2 can only be a presentation of the last 2 s of the clip or a novel video). Next, we performed a positive control where we compared the same segment of video between two different clips. We used the firing rates from 0.1-2.1 s after video presentation during both the clip 1 and clip 3 periods. Finally, we performed a positive control where we compared firing rates to different segments within the same clip, even though different videos were shown in these segments. For this we correlated the firing rates between 0.1-2.1 and 2.1-4.1 s after the presentation of clip 3.

Acknowledgements
We thank Andrea Shang, Ellen Wang, and Joyce Ho for expert animal care and Yuji Naya for suggestions in task development. This research was supported by the National Institutes of Health (R01-MH084964) and the Simons Collaboration on the Global Brain (542997).

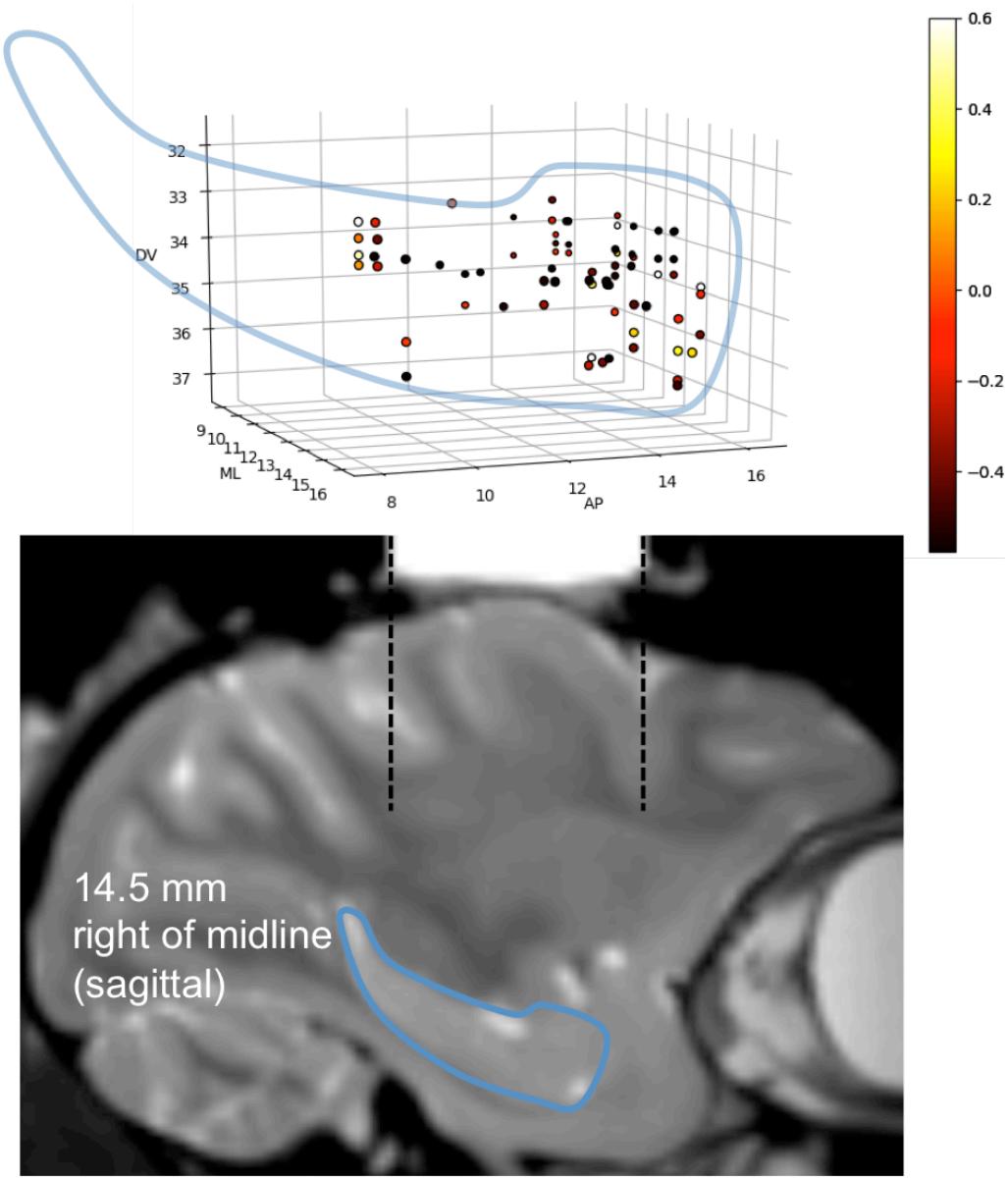
Author Contributions
JJS and WAS designed the experiment and wrote the paper. JJS collected the data and designed and performed the analyses.

Data availability
Full data, or data pre-processed for a specific analysis, is available upon request.

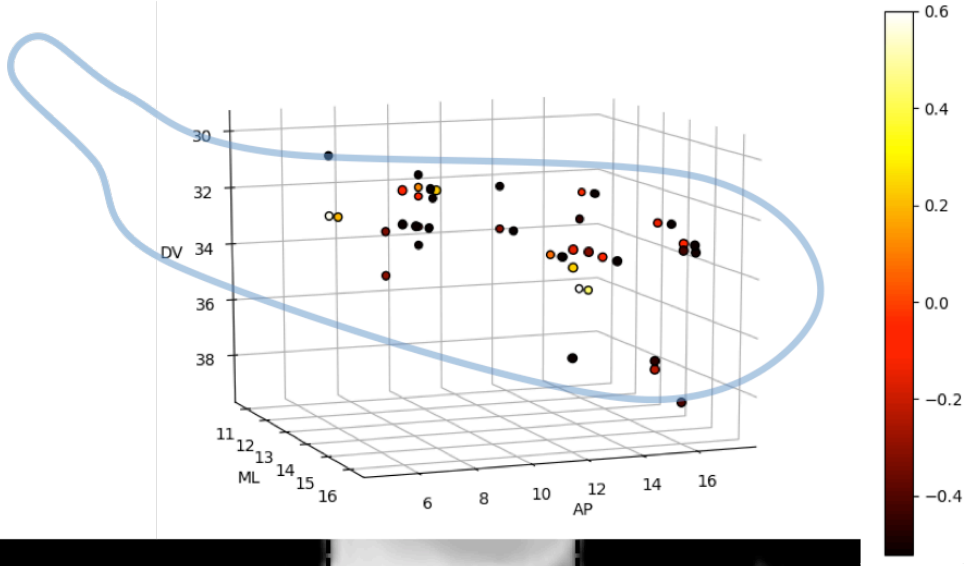
Code availability
Analysis code (in Matlab) and pre-processed data to reproduce analyses and figures are available on request.

Figure S1 – Hippocampal recording locations for each animal. Each point represents the recording site of a neuron. **Top:** The outline of the hippocampus, reproduced from the MRI image at bottom, gives an estimate of the relative locations within the hippocampus of each animal. The colors of each point are the weights from the classifier for Subclip 1 (Fig. 5A), and the sizes indicate the ML position with smallest being closest to midline (min=9.5, max=16). Note the heterogeneity of the weights across the hippocampus, with neurons showing strong repeat v. novel classification throughout the anterior-posterior extent in both monkeys. The top-6 weights (5 from monkey I, 1 from monkey W) were each set to 0.6 to increase the dynamic range of the colorscale. **Bottom:** Sagittal MRI images at indicated distance from midline with hippocampus outlined in blue. Dashed lines show the extent of the recording chambers. All electrodes were inserted through gridholes parallel to these lines. AP=anteroposterior. ML=mediolateral. DV=dorsoventral.

Monkey I:



Monkey W:


DV
30
32
34
36
38
ML
11
12
13
14
15
16
AP
6
8
10
12
14
16
0.6
0.4
0.2
0.0
−0.2
−0.4

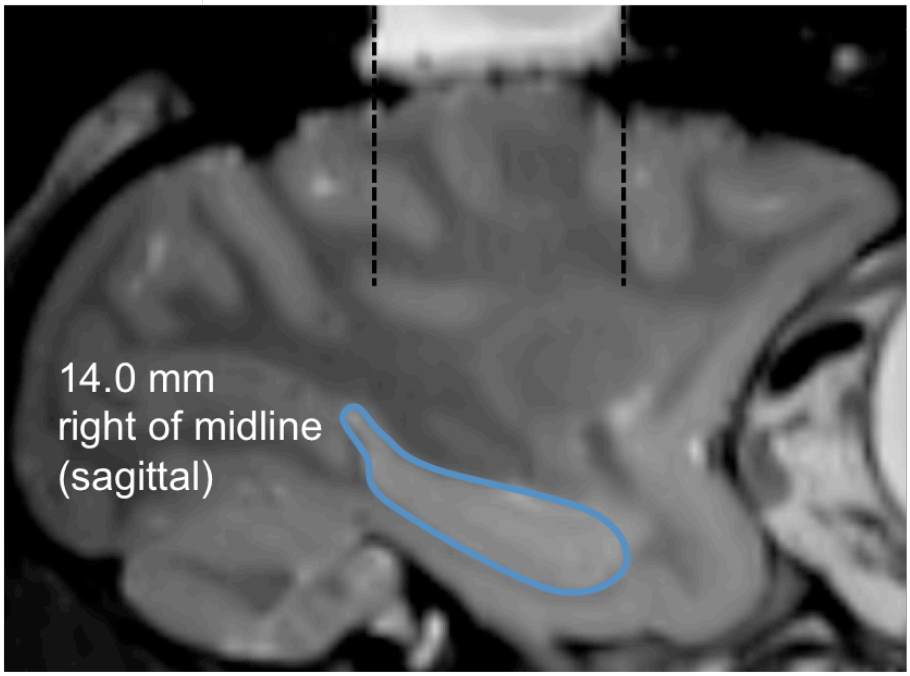
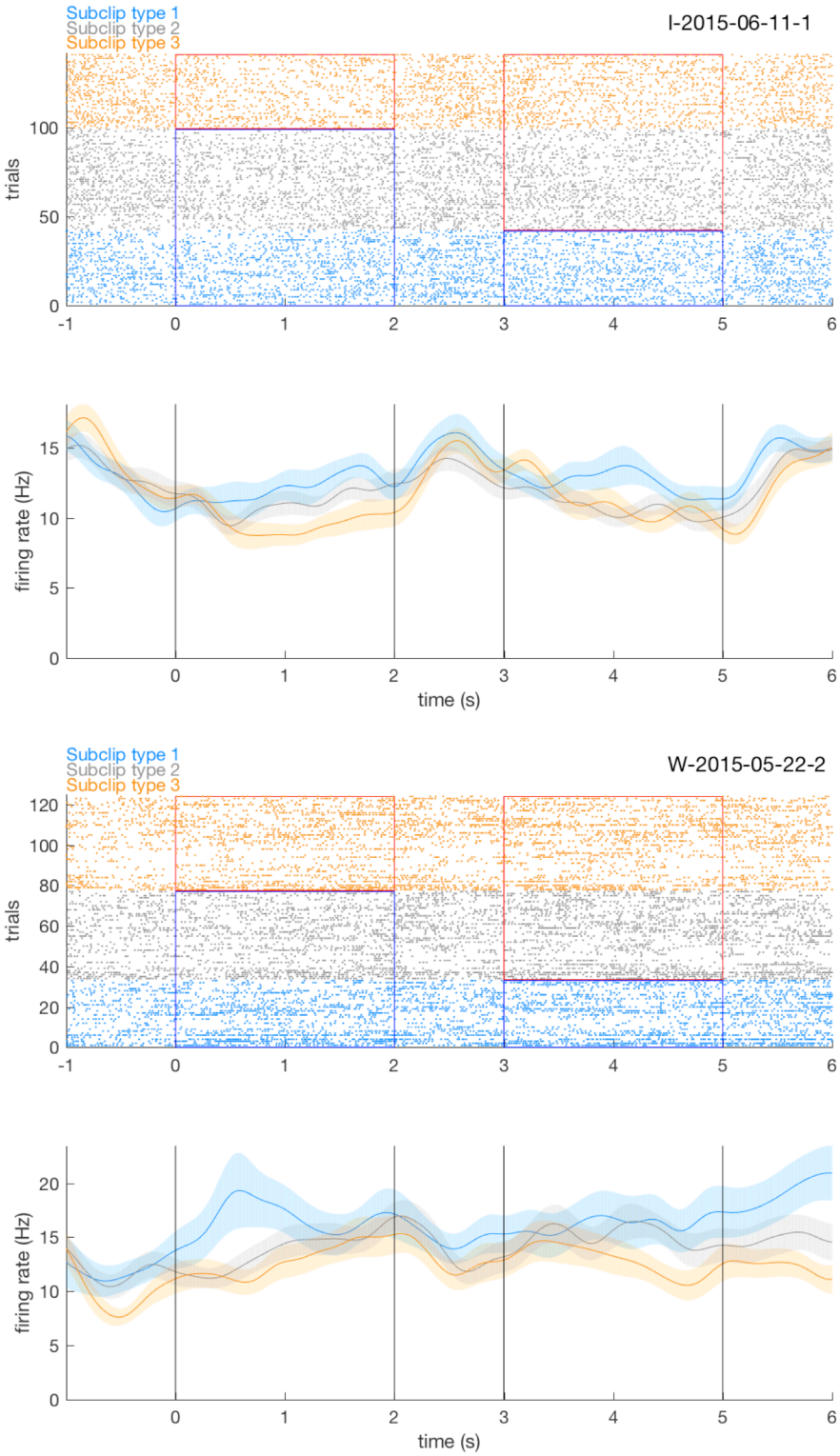
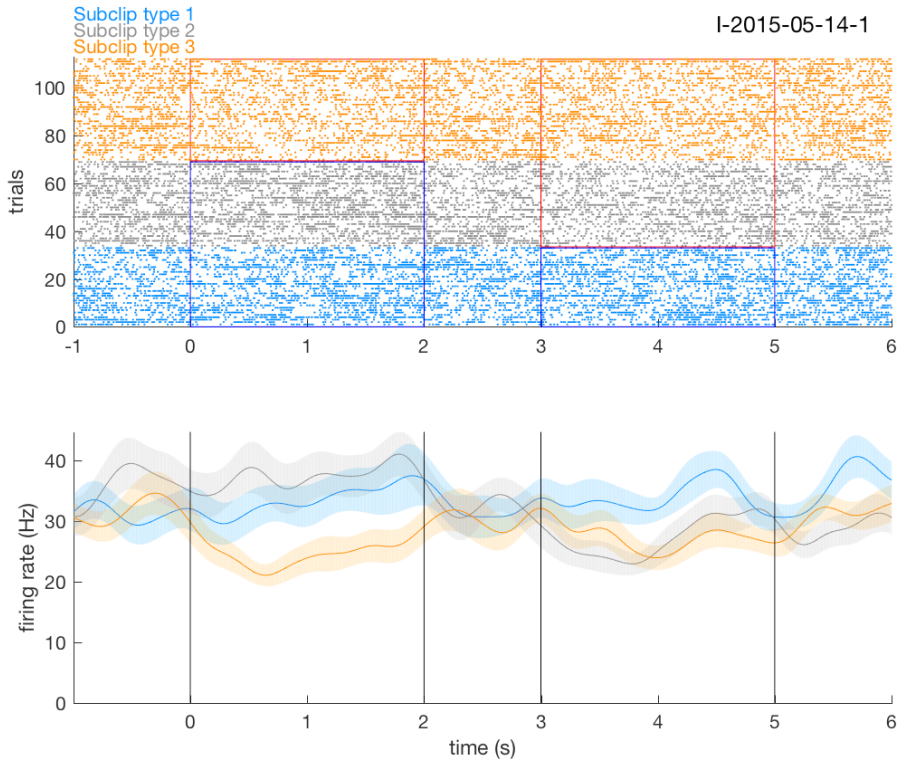

14.0 mm
right of midline
(sagittal)

Figure S2 – Additional examples of individual neurons. Conventions as in Fig. 4. These neurons were the 8th, 6th, 15th, and 22nd most positive weights in the classifier for the subclip 1 period and the 9th, 22nd , 4th, and 14th most positive weights in the classifier for the subclip 2 period (Fig. 2).




Subclip type 1
Subclip type 2
Subclip type 3
I-2015-05-14-1
trials
firing rate (Hz)
time (s)

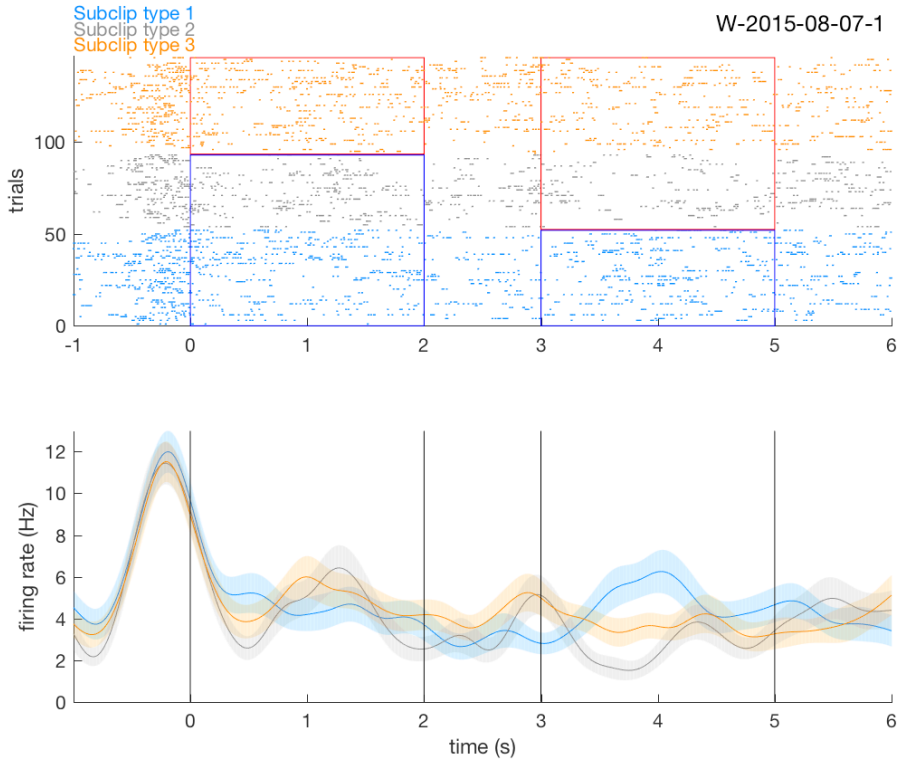
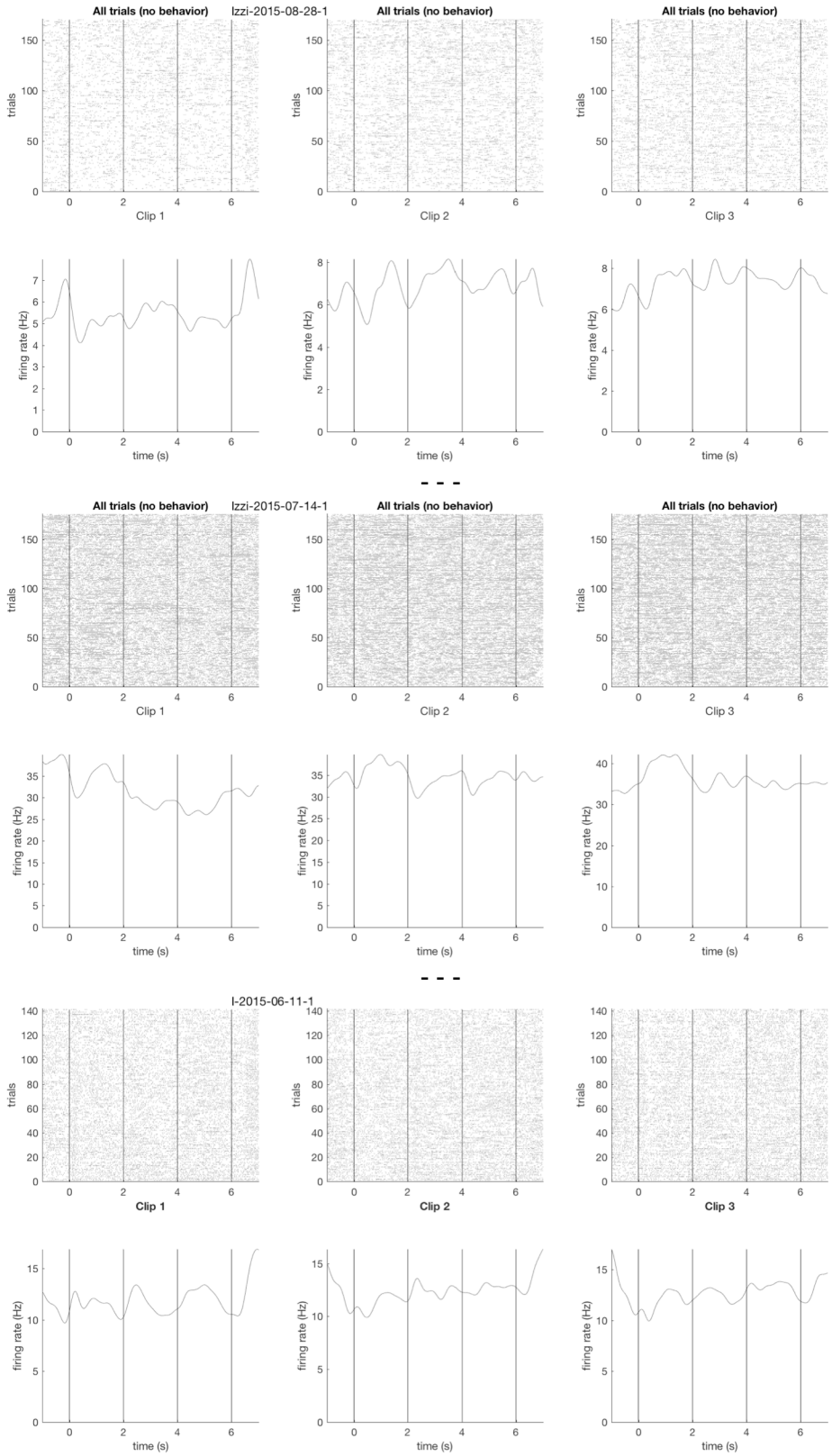
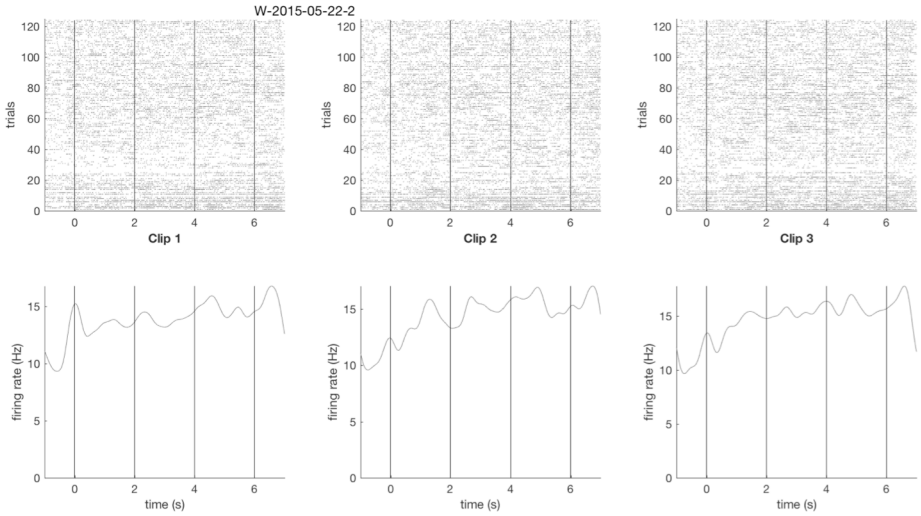
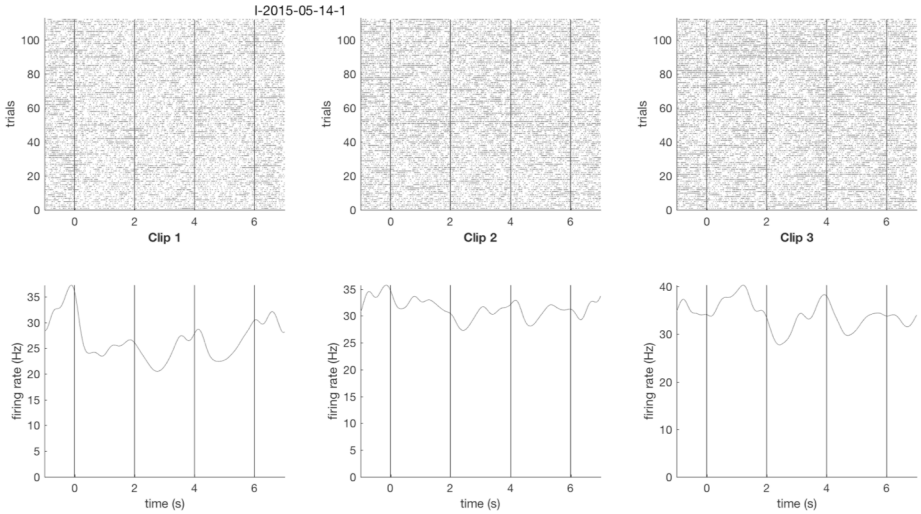

Subclip type 1
Subclip type 2
Subclip type 3
W-2015-08-07-1
trials
firing rate (Hz)
time (s)

Figure S3 – Rasters and PDFs of clips 1-3 for example neurons in Fig. 4 and Fig. S2.




W-2015-05-22-2
trials
Clip 1
Clip 2
Clip 3
firing rate (Hz)
time (s)

- - -


I-2015-05-14-1
trials
Clip 1
Clip 2
Clip 3
firing rate (Hz)
time (s)

- - -

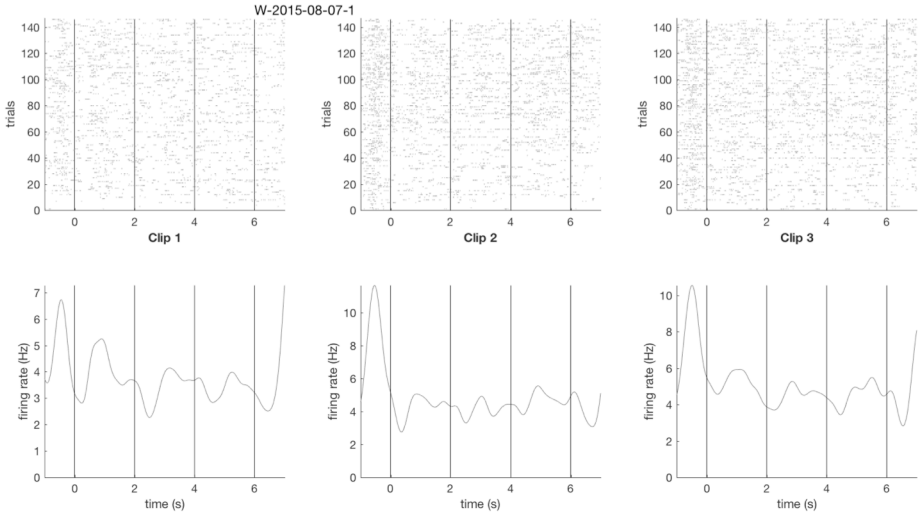
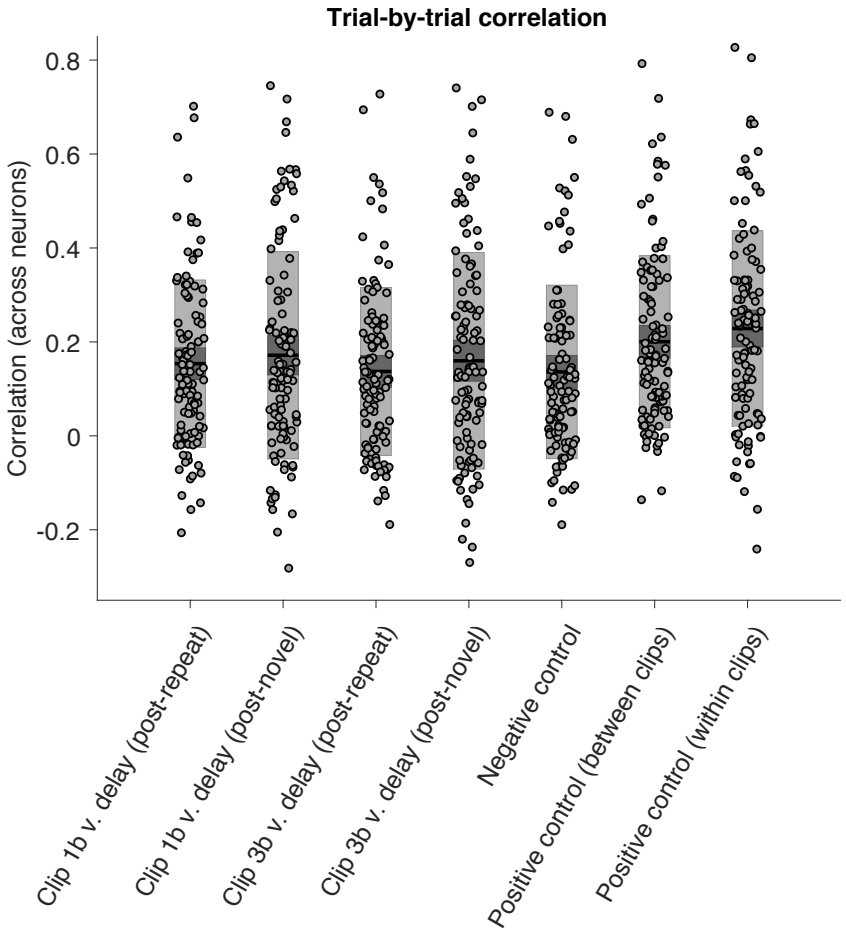

W-2015-08-07-1
trials
Clip 1
Clip 2
Clip 3
firing rate (Hz)
time (s)

Figure S4 – Full distributions of correlations across neurons during key periods. The same plots as Figure 6, but with the full distributions of neurons shown. Shaded boxes are the 95% confidence intervals (dark gray) and the standard deviations (light gray).



References

Bakker, A., Kirwan, C.B., Miller, M., and Stark, C.E.L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. Science *319*, 1640-1642.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., and Norman, K.A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. Neuron *95*, 709-721 e705.

Baraduc, P., Duhamel, J.R., and Wirth, S. (2019). Schema cells in the macaque hippocampus. Science *363*, 635-639.

Barron, H.C., Garvert, M.M., and Behrens, T.E. (2016). Repetition suppression: a means to index neural representations using BOLD? Philosophical transactions of the Royal Society of London Series B, Biological sciences *371*.

Baxter, M.G. (2009). Involvement of medial temporal lobe structures in memory and perception. Neuron *61*, 667-677.

Ben-Yakov, A., Eshel, N., and Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. Journal of experimental psychology General *142*, 1255-1263.

Bird, C.M. (2017). The role of the hippocampus in recognition memory. Cortex; a journal devoted to the study of the nervous system and behavior *93*, 155-165.

Brown, M.W., Wilson, F.A., and Riches, I.P. (1987). Neuronal evidence that inferomedial temporal cortex is more important than hippocampus in certain processes underlying recognition memory. Brain Res *409*, 158-162.

Brown, M.W., and Xiang, J.Z. (1998). Recognition memory: neuronal substrates of the judgement of prior occurrence. Prog Neurobiol *55*, 149-189.

Chadwick, M.J., Hassabis, D., Weiskopf, N., and Maguire, E.A. (2010). Decoding individual episodic memory traces in the human hippocampus. Curr Biol *20*, 544-547.

Dimsdale-Zucker, H.R., Ritchey, M., Ekstrom, A.D., Yonelinas, A.P., and Ranganath, C. (2018). CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. Nat Commun *9*, 294.

Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., and Lipton, P. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. Neurosci Biobehav Rev *36*, 1597-1608.

Ezzyat, Y., and Davachi, L. (2011). What constitutes an episode in episodic memory? Psychol Sci *22*, 243-252.

Fried, I., MacDonald, K.A., and Wilson, C.L. (1997). Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. Neuron *18*, 753-765.

Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., and Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. Science *322*, 96-101.

Heit, G., Smith, M.E., and Halgren, E. (1990). Neuronal activity in the human medial temporal lobe during recognition memory. Brain *113 (Pt 4)*, 1093-1112.

Horner, A.J., Bisby, J.A., Bush, D., Lin, W.J., and Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. Nat Commun *6*, 7462.

Jutras, M.J., and Buffalo, E.A. (2010). Recognition memory signals in the macaque hippocampus. P Natl Acad Sci USA *107*, 401-406.

Kiani, R., Cueva, C.J., Reppas, J.B., and Newsome, W.T. (2014). Dynamics of Neural Population Responses in Prefrontal Cortex Indicate Changes of Mind on Single Trials. Curr Biol *24*, 1542-1547.

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron *93*, 480-490.

Kraus, Benjamin J., Robinson Ii, Robert J., White, John A., Eichenbaum, H., and Hasselmo, Michael E. (2013). Hippocampal "Time Cells": Time versus Path Integration. Neuron *78*, 1090-1101.

Lacy, J.W., Yassa, M.A., Stark, S.M., Muftuler, L.T., and Stark, C.E.L. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. Learn Memory *18*, 15-18.

Lenat, D.B., Guha, R.V., Pittman, K., Pratt, D., and Shepherd, M. (1990). Cyc: toward programs with common sense. Commun ACM *33*, 30-49.

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Ranganath, C., and Redish, A.D. (2017). Viewpoints: how the hippocampus contributes to memory, navigation and cognition. Nat Neurosci *20*, 1434-1447.

Marcus, G. (2018). Deep Learning: A Critical Appraisal. arXiv e-prints *arXiv:1801.00631*.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods *164*, 177-190.

Merkow, M.B., Burke, J.F., and Kahana, M.J. (2015). The human hippocampus contributes to both the recollection and familiarity components of recognition memory. P Natl Acad Sci USA *Early Edition*.

Meyer, T., and Rust, N.C. (2018). Single-exposure visual memory judgments are reflected in inferotemporal cortex. Elife *7*.

Miller, E.K., Li, L., and Desimone, R. (1991). A Neural Mechanism for Working and Recognition Memory in Inferior Temporal Cortex. Science *254*, 1377-1379.

Minsky, M. (1986). The society of mind (New York: Simon and Schuster).

Naya, Y., and Suzuki, W.A. (2011). Integrating What and When Across the Primate Medial Temporal Lobe. Science *333*, 773-776.

Quiroga, R.Q. (2012). Concept cells: the building blocks of declarative memory functions. Nat Rev Neurosci *13*, 587-597.

Quiroga, R.Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput *16*, 1661-1687.

Riches, I.P., Wilson, F.A., and Brown, M.W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. J Neurosci *11*, 1763-1779.

Rolls, E.T., Cahusac, P.M., Feigenbaum, J.D., and Miyashita, Y. (1993). Responses of single neurons in the hippocampus of the macaque related to recognition memory. Experimental brain research *93*, 299-306.

Rolls, E.T., and Treves, A. (2011). The neuronal encoding of information in the brain. Prog Neurobiol *95*, 448-490.

Rutishauser, U., Mamelak, A.N., and Schuman, E.M. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. Neuron *49*, 805-813.

Rutishauser, U., Schuman, E.M., and Mamelak, A.N. (2008). Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. Proc Natl Acad Sci U S A *105*, 329-334.

Sakon, J.J., Naya, Y., Wirth, S.,Suzuki,W.A. (2014). Context-dependent incremental timing cells in the primate hippocampus. P Natl Acad Sci USA *(in press)*.

Sakon, J.J., and Suzuki, W.A. (2019). A neural signature of pattern separation in the monkey hippocampus. Proc Natl Acad Sci U S A *116*, 9634-9643.

Shadlen, M.N., and Shohamy, D. (2016). Decision Making and Sequential Sampling from Memory. Neuron *90*, 927-939.

Spelke, E.S. (1985). Preferential looking methods as tools for the study of cognition in infancy. In Measurement of audition and vision in the first year of life, G.G.N. Krasnegor, ed. (Norwood, NJ: Ablex), pp. 323-363.

Stokes, M.G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. Neuron *78*, 364-375.

Suzuki, W.A., Miller, E.K., and Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. J Neurophysiol *78*, 1062-1081.

Viskontas, I.V., Knowlton, B.J., Steinmetz, P.N., and Fried, I. (2006). Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. J Cogn Neurosci *18*, 1654-1662.

Voss, J.L., and Paller, K.A. (2008). Brain substrates of implicit and explicit memory: the importance of concurrently acquired neural signals of both memory types. Neuropsychologia *46*, 3021-3029.

Wais, P.E., Squire, L.R., and Wixted, J.T. (2010). In Search of Recollection and Familiarity Signals in the Hippocampus. J Cognitive Neurosci *22*, 109-123.

Wais, P.E., Wixted, J.T., Hopkins, R.O., and Squire, L.R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. Neuron *49*, 459-466.

Wirth, S., Avsar, E., Chiu, C.C., Sharma, V., Smith, A.C., Brown, E., and Suzuki, W.A. (2009). Trial Outcome and Associative Learning Signals in the Monkey Hippocampus. Neuron *61*, 930-940.

Wirth, S., Yanike, M., Frank, L.M., Smith, A.C., Brown, E.N., and Suzuki, W.A. (2003). Single neurons in the monkey hippocampus and learning of new associations. Science *300*, 1578-1581.

Xiang, J.Z., and Brown, M.W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. Neuropharmacology *37*, 657-676.

Yonelinas, A.P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. J Mem Lang *46*, 441-517.