PNAS
www.pnas.org

Supplementary Information for

Working memory capacity is enhanced by distributed prefrontal activation and invariant temporal dynamics

Hua Tang, Xue-Lian Qi, Mitchell R. Riley, and Christos Constantinidis

Email: cconstan@wakehealth.edu

This PDF file includes:

Supplementary text
Figs. S1 to S15
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1818278116

Supplementary Information Text

Experimental Procedures

Surgery and neurophysiology. Two, 20-mm-diameter recording cylinders were implanted over the dlPFC and the PPC in each monkey (Fig. 2A and Fig. S11A). Extracellular activity of single units was recorded from areas 8a and 46 of dlPFC and areas 7a and the lateral intraparietal area (LIP) of the PPC. The anatomic location of electrode penetrations was determined on the basis of MR imaging. Recordings were obtained with arrays of two to four microelectrodes in each cylinder. Epoxylite-coated Tungsten electrodes were used with a 250 μm diameter and 1-4 MΩ impedance at 1 kHz (FHC, Bowdoin, ME). The electrical signal from each electrode was amplified, bandpass filtered between 500 Hz and 8 kHz, and recorded with a modular data acquisition system at 25-μs resolution (APM system; FHC, Bowdoin, ME). Waveforms that exceeded a user-defined threshold were sampled at 25 μs resolution, digitized, and stored for off-line analysis. Neurons were sampled in an unbiased fashion, collecting data from all units isolated from our electrodes, with no regard to response properties of a neuron being isolated.

Behavioral tasks. The monkeys faced a computer monitor 60 cm away in a dark room with their head fixed, as described in detail previously (1). Eye position was sampled at 240 Hz, digitized, and recorded with an infrared eye position tracking system (model RK-716; ISCAN, Burlington, MA). The visual stimulus presentation and behavior monitoring were controlled by in-house software (2) implemented in the MATLAB computational environment (Mathworks, Natick, MA), using the Psychophysics Toolbox (3). The system achieved a $< 0.3°$ resolution around the center of vision. Monkeys were required to maintain their gaze on a 2° fixation target throughout a trial; breaks in fixation aborted the trial.

The monkeys were trained in a Match/Nonmatch task requiring them to observe a display with multiple stimuli, to maintain their location in memory, and to judge if a second display contained items at identical locations or not. The two animals were tested with slightly different variations of the task. For one monkey, after the two displays were presented in sequence, two colored squared appeared, one of which was green and the other one blue. The monkey was required to saccade to the green square if the two stimulus displays had been identical, and to the blue square, otherwise. The monkey received liquid reward after staying in the correct choice target for 100 ms. The second animal was trained to pull a lever to initiate the trial. After the second display presentation, the monkey had to continue holding the lever for 500 ms if the stimulus displays had been identical, and to release the lever within 500 ms if they were not. The monkey received a liquid reward for a correct response. The trial was immediately aborted if the lever was released at any other time during the trial, or if the monkey broke fixation at any point prior to the lever release. During recording sessions, we typically used 30 different possible displays, containing one to five stimuli (6 different arrangements of each number of stimuli), pseudo-randomly interleaved. These changed from day to day. To ensure that monkeys did not abort the most difficult trial types and complete just the easiest ones, a block of all possible trial types needed to be completed correctly, presented in pseudo-random sequence, before trials of the next block were presented. Additional details about the task and performance of the animals for different types of displays have been described elsewhere (4).

Behavioral Data Analysis. Our primary measure of performance was the percentage of trials that the monkey correctly identified as a Match or as a Nonmatch, in daily sessions. This was plotted as a function of stimuli in each display. We also estimated capacity based on the Pashler Formula (5):

$$K = \frac{h - fa}{1 - fa} \times N$$

where *h* represents the hit rate, *fa* the false alarm rate, and *N* the set size. In each session, we determined the average K value obtained for displays of four and five stimuli, and then relied on a median split to distinguish between low- and high-performance sessions.

Neural Data Analysis. All data analysis was implemented with the MATLAB computational environment. Recorded spike waveforms were sorted into separate units using an automated cluster analysis relying on the KlustaKwik algorithm (6), which relied on principal component analysis of the waveforms.

Neurons with selectivity between the stimulus arrangements (typically 30 different displays used in a single session) were identified as those that exhibited a significantly different firing rate averaged during the cue period, or the delay period, across all stimuli used to test them (1-way ANOVA, $p<0.05$). Most analyses relied on these neurons, and on data from correct trials. The effect of different numbers of stimuli on firing rate was evaluated by computing the average firing rate during the fixation, cue, and delay period, and using a repeated-measures ANOVA to compare responses of the same neurons to displays with different numbers of stimuli.

Responses during the sample (second stimulus presentation) period were used to determine the Match/Nonmatch selectivity of each neuron. As different prefrontal neurons have an overall preference for either Match or Nonmatch stimuli, we relied on the absolute value of |Match-Nonmatch| responses. We averaged this measure across neurons, and across displays of different number of stimuli.

Demixed Principal Component Analysis (dPCA) was performed as we have described elsewhere (7). The method treats the responses of each neuron to one type of stimulus condition as one dimension, and then performs dimensionality reduction to determine components that correspond to stimulus and task variables. As different stimulus displays were used each day (to avoid the monkey memorizing them), we rank-ordered stimulus displays depending on firing rate during the delay period, grouped them into six groups, and then treated trials of each of groups as a single stimulus, across neurons.

Analysis of error trials relied only on trials in which the monkey did not abort the trial prematurely, but made the wrong selection when the choice targets appeared (indicating a Match, when the second display was a Nonmatch, and vice versa). For some analyses firing rate in error trials were compared with that of correct; in others, correct and error trials were pooled together and were compared with correct trials alone.

A receiver operating characteristic (ROC) analysis was used to determine the ability of neurons to discriminate between different displays. ROC values compared the distribution of firing rates for the best and worst display (based on mean cue responses) having the same number of stimuli. A single ROC value was determined for each neuron by averaging values across displays of 1-5 stimuli. ROC values were computed separately for low- and high-performance sessions in 250 ms bins, stepped every 10 ms.

We computed the autocorrelation of firing rate in the fixation period, a measure of the neuron's intrinsic timescale that quantifies the stability of firing dynamics (8). We also computed autocorrelation of firing rate across the trial, as a measure of how stereotypical responses were. For this analysis, firing rates were grouped across all displays of equal number of stimuli and then the Pearson correlation coefficient was computed between the firing rate of a neuron at the first 500 ms of the fixation period, and every subsequent 500 ms interval. These values were then averaged across neurons.

The Fano factor of a neuron's spike counts (defined as the variance divided by the mean) was estimated across the length of the trial, as described previously (9, 10). Data for each neuron were treated separately. Spike counts were computed in a 50-ms sliding window moving in 20-ms steps across the length of the trial, separately for each stimulus condition. We computed the variance and mean of the spike count of each neuron across trials and performed a regression of the variance to the mean (9). This slope of this regression represents the Fano factor reported

here. Fano factors were determined separately for each neuron, and then averaged across neurons. To compare two conditions (sessions of high vs. low performance), we created distributions of firing rates by all available neurons at each time point, and only included neurons with matching firing rates, excluding neurons whose (too low or too high) firing rate could not be matched.

Supplementary Results

Analysis of aborted trials. Results of performance in the main text were based on correct trials, divided by the number of correct plus erroneous, completed trials. We also analyzed non-completed trials, in which the monkey broke fixation before the end of the trial. We focused particularly on trials terminated during the first stimulus presentation, delay period, or sample presentation. We refer to these as aborted trials. The rate of aborted trials was dependent of number of stimuli; the monkey was more likely to abort a trial with more stimuli (Fig. S2). Furthermore, this rate increased in high-performance sessions, possibly due to awareness of internal state in a particular trial. The monkeys were highly trained in the single-stimulus version of the task, before the training with multiple stimuli began, and the rate of aborted trials was likely influenced by the confidence of the monkey in completing the trial.
We should point out that the performance rate we report was not affected by aborted trials: First, the rate of performance was calculated based on completed trials alone, and the rate of aborted trials did not influence it. Second, when a trial was aborted, this was repeated later in the block, until all different stimulus sets had been presented once, before another block of stimuli could be presented. Thus, the monkeys could not improve their performance rate simply by aborting the most difficult trial types.

Comparison of neural activity in Early and Late sessions. Our analysis in the main text relied on a median split based on sessions where the monkey achieved higher or lower performance. Very similar conclusions were reached when we broke down sessions entirely chronologically, based on a median split on recording date. Fifty-five selective neurons were recorded from early and 56 neurons in late sessions. The results are shown in Fig. S6. Late sessions were characterized by lower firing rate in the baseline fixation period (Fig. S6C), as was the case of sessions selected based on high performance. Cue activity was also diminished (Fig. S6D) though the delay period firing rate was comparable in the early- and late-sessions (Fig. S6E). As was the case of sessions selected based on high performance, activity in late sessions became highly stereotypical, and subtracting fixation activity from delay period resulted in very low variability (Fig. S6H).

We also considered simultaneously the effects of session number and performance by dividing sessions in four groups, considering both early vs. late and high vs. low performance (Fig. S12). The effects of cumulative training appeared to be important independent of performance in this analysis. A further decline in baseline firing rate and more stereotypical time course of activation was evident in late sessions in which the monkey achieved low performance (Fig. S12C) relative to early sessions where the monkey achieved high performance (Fig. S12B), even though the average level of performance in these two groups was essentially identical (Fig. S12E).

Posterior parietal recordings. Recordings were performed in the posterior parietal cortex (PPC), including areas 7a and LIP (Fig. S11A), a cortical region providing direct afferent input to the dorsolateral prefrontal cortex (11). A total of 250 neurons were recorded from monkey EL, in an additional series of recordings, interleaved with the prefrontal recordings. Of those, 72 neurons exhibited selectivity for the spatial locations of the stimuli during the cue or delay period (1-way ANOVA, $p<0.05$). A total of 162 neurons were recorded in low-performance sessions based on a median split depending on performance. Of those, 33 (20%) were selective for the stimulus pattern. A total of 88 neurons were recorded in high performance sessions. The percentage of neurons selective for stimulus pattern increased substantially to 44% (39/88 neurons), a significant difference (χ^2 test, $p = 6.51 \times 10^{-5}$).

Activity changes in the posterior parietal cortex mirrored that of the prefrontal cortex (Fig. S11B,C). Activity in PPC also declined in sessions of high performance over lower performance (Fig. S11D-G). A 2-way ANOVA with factors the number of stimuli and performance revealed a significant decrease in the fixation period ($F_{1,350}$=16.55, p = 5.85 x10^{-5}), the cue period ($F_{1,350}$=9.92, p = 1.77 x10^{-3}), the delay period ($F_{1,350}$=21.35, p = 5.38x10^{-6}) and sample period ($F_{1,350}$=7.80, p = 5.51 x10^{-3}). There was no significant difference when we compared the evoked activity for the cue ($F_{1,350}$=1.63, p = 0.202), delay ($F_{1,350}$=3.30, p = 0.070) or sample period ($F_{1,350}$=1.34, p = 0.249), relative to the baseline (Fig. S11H-J). Invariant temporal dynamics were also predictive of correct and error trials (Fig. S11L). However, the change in temporal dynamics was less pronounced in the posterior parietal cortex, and the time course in high-performance sessions not as stereotypical as that seen in the prefrontal cortex (Fig. S11K).

Comparison of Low- vs. High- Performance sessions based on all neurons. We replicated the analysis of Fig. 4, relying not only on neurons that were selective for the different stimuli, but on all neurons recorded from any of our electrodes. A total of 203 neurons were recorded in the low- and 102 neurons in the high-performance sessions. Results are shown in Fig. S13. All changes in firing rate between low- and high-performance sessions moved in the same direction as those reported for the selective neurons, in the main text. Most notably, a decrease in baseline firing rate was evident in sessions of high- compared to sessions of low-performance. The only departure from the analysis shown in Fig. 4 was that, since more neurons became responsive to stimuli in the high-performance sessions, the average responses of all neurons during the earliest time point of the cue were higher for the high-performance sessions (notice height of cue period peaks in Fig. S13A,B). Even in this case however, integrated over the entire duration of the cue period, the average cue response in the high-performance sessions ended up being no higher than of low-performance sessions, since firing rate rose from and returned to a lower baseline (Fig. S13D).

Fano factor of neuronal responses. We examined the trial-to-trial variability of neuronal responses by estimating the Fano factor (variance divided by mean) of spike counts, and tested whether variability differed between low and high performance sessions. Fano factor values were generally higher in low-performance sessions, however it is well understood that Fano factors increase with firing rates (12, 13), and firing rates declined after training. Therefore, we performed all comparisons to firing-rate matched groups of neurons. As the time course of activity was highly dynamic, even in the fixation period, we additionally performed this comparison separately at each time point, each time comparing neurons with firing rate matched in the time point under study (Fig. S14). The analysis revealed a slight decrease in Fano factor in the fixation period in high performance sessions.

Anatomical localization. Recordings were obtained from the same regions of the prefrontal cortex in the low- and high-performance sessions in the two animals. In some instances, the same tracks were revisited, and neurons from the same locations were used for the low- and high-performance session comparison. The complete map of penetrations is shown in Fig. S15.

Supplementary Discussion

Nature of performance improvement. Continued practice of the behavioral task resulted in improved performance during the time course of the experiments. There are a number of possible sources of behavioral improvement that could account for improvement. First, the monkeys may have simply understood the basic rules and requirements of the task better (i.e. that a match stimulus required a response towards the green choice target). This possibility is unlikely, as the monkeys were well trained with the single-stimulus version of the task prior to beginning practice of the multi-stimulus displays and no benefits in performance was observed for stimulus sets with 1-3 stimuli (Fig. 1G). A related possibility would be that improved performance was achieved through higher levels of attention and vigilance in higher-performance sessions, which would suggest lower engagement with the task and higher rates of aborted trials in early sessions. This was clearly not the case, either. The monkeys produced higher rates of aborted trials in high-performance sessions (Fig. S2). Alternatively, the monkeys may have exhibited improved efficiency with respect to management of working memory content or general improvement in information processing, and network stability. It appears likely that such a benefit was indeed realized. Correct trials were characterized by lower firing rates and stereotypical ramping of firing rate already present in the fixation period of the trial. Their pattern of aborted trials suggests that the monkeys had awareness of this internal state and aborted trials that were likely to result in errors. Finally, the monkeys may have improved working memory capacity, irrespective of any improvements in task management efficiency. This type of improvement also seems likely, as performance improvements primarily involved displays with the largest numbers of stimuli (Fig. 1G). In contrast, the increase of numbers of aborted trials in high- relative to low-performance trials was essentially uniform across set sizes, as would be expected by an efficiency mechanism which operates independently of number of stimuli maintained in memory.

Imaging studies of working memory capacity

Some human studies have suggested increases in activity after training that improves working memory performance (14-19). These increases may reflect an increase in the number of neurons activated during working memory and/or a higher firing rate. Indeed, previous studies of training in (single-stimulus) working memory tasks have shown a general increase in the number of activated prefrontal neurons and their average firing rate after learning to perform the task for the first time (1), and as performance improves in that task (20). However, other human fMRI studies have documented decreases in activity (21-24). These results have been interpreted as suggestive of increases in efficiency, which could be due to strategic allocation of fewer neural resources to achieve the same goal, or effective reduction of the load of the task by inventing shortcuts akin to "chunking". The underlying neural processes of efficiency have been hitherto unknown.

The relationship between changes in neural activity such the ones we observed and BOLD activation is complex and in many cases non-linear (25). In our study, more neurons were responsive after practice, but an overall decrease in baseline prefrontal activity was observed among neurons that were selective for the stimuli. Averaging activity across the duration of the trial and across multiple task conditions might yield an overall increase or decrease in activity, depending on the specifics of each experiment. The effects we observed were not exclusive to the prefrontal cortex; a separate series of recordings in the posterior parietal cortex revealed similar recruitment of more neurons, and lower baseline discharge rate. Indeed, imaging studies have also identified parallel changes in prefrontal and posterior parietal activity with training (19).

Computational models of capacity limitation

The schematic of Figure S1 depicts the neuronal mechanisms that may cause errors in performance during working memory tasks with multiple items. Drift of the peak (bump) of

activity in the network with time (Fig. S1A); decay of information corresponding to one of the items (Fig. S1B); and merging of two items into one (Fig. S1C). Computational models provide evidence for such mechanisms (26-28). In particular, experimental evidence suggests that drift best accounts for patterns of neural activity and behavior observed in a single-item working memory task in non-human primates (26) and psychophysical behavior of multi-item working memory tasks in humans (29). On the other hand, these studies provided evidence against an alternative model that that working memory fades in time due to a bump of activity of increasing variance during the delay period. Such a mechanism seems intuitively plausible as the variance of behavioral errors increases with time (30), however modeling and experimental results suggest that a broadening of the bump of neural activity is unlikely (26, 29).

Supplementary Figures

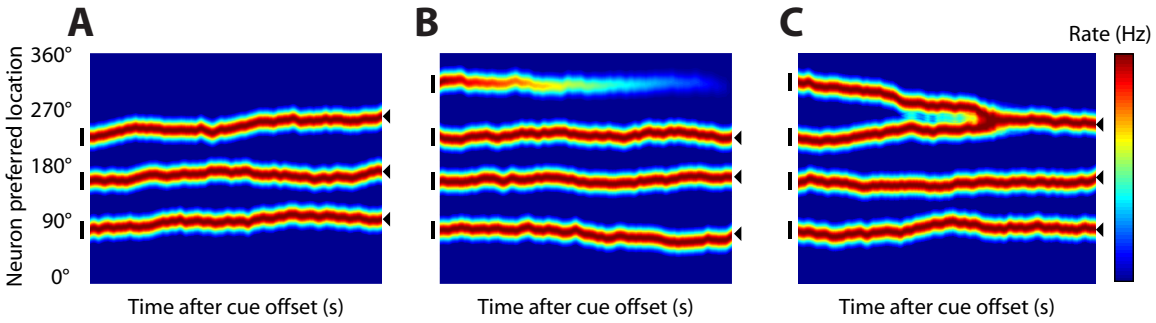

Fig. S1. Schematic of bump attractor model. (A) Firing rate of different neurons, plotted in the abscissa based on their preferred location varying from 0° to 360°, is represented as a function of time after three stimuli have appeared at different locations (vertical bars). Three bumps of activity persist in the delay period after the cue offset, although the peaks may drift with time, resulting in imprecise recall of the stimulus locations at the end of the delay period (triangles). (B) Activity in the same network when four stimuli have been presented, exceeding the capacity of the system; decay of activity related to the top stimulus results in recall of three stimuli. (C) Activity in the network, after drift causes the bumps of two stimuli to merge with each other, again, leading to loss of one stimulus information.

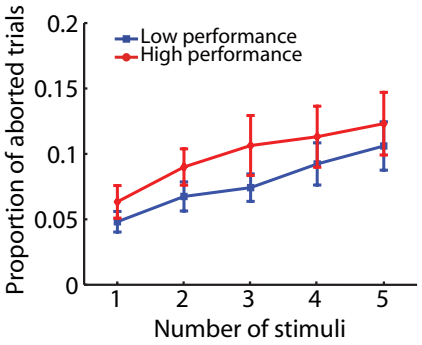

Fig. S2. Aborted trials. The proportion of trials that were aborted is plotted as a function of number of stimuli, in low- and high-performance sessions. In each case, the ratio of aborted trials is divided by the sum of aborted and correct trials to produce the proportion being plotted. Data from two monkeys, n=111. Bars represented standard error of the mean across sessions.

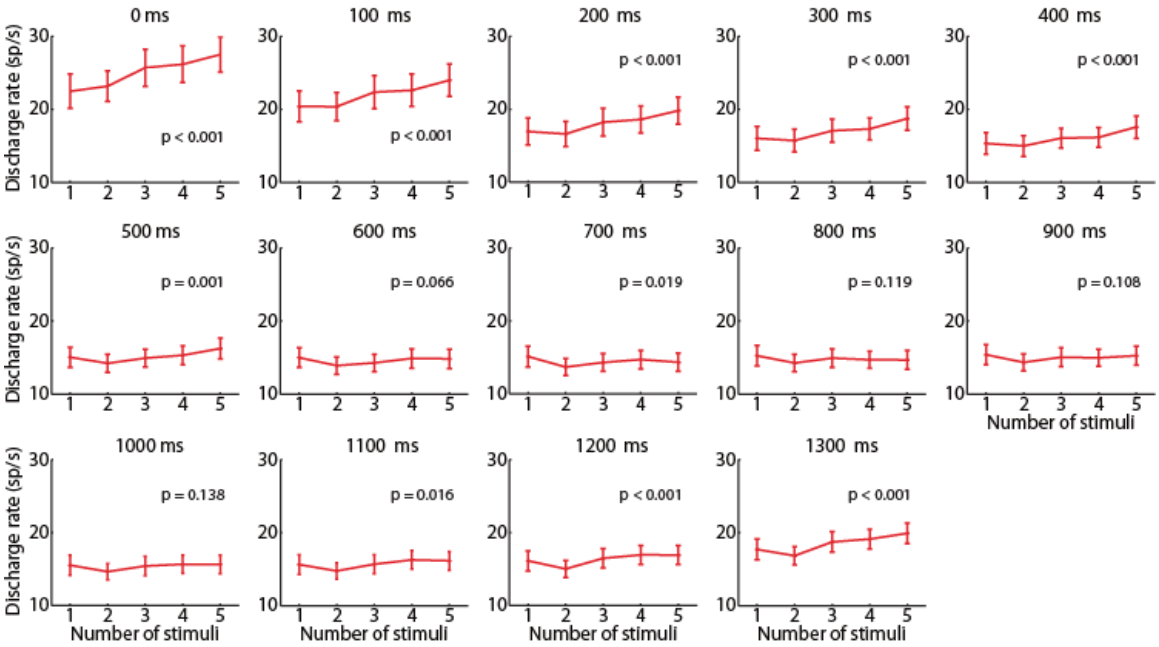

Fig. S3. Time course of activity as a function of number of stimuli. Mean firing rate for displays of different number of stimuli is depicted at successive time points during the trial, from the onset of the cue, until the end of the delay period, for the same population of neurons. Each panel represents activity averaged in a 250 ms bin; successive panels represent a step of 100 ms. Time point 0 corresponds to the onset of cue period; time point 500, the onset of the delay period. Data from two monkeys, n=111 neurons.

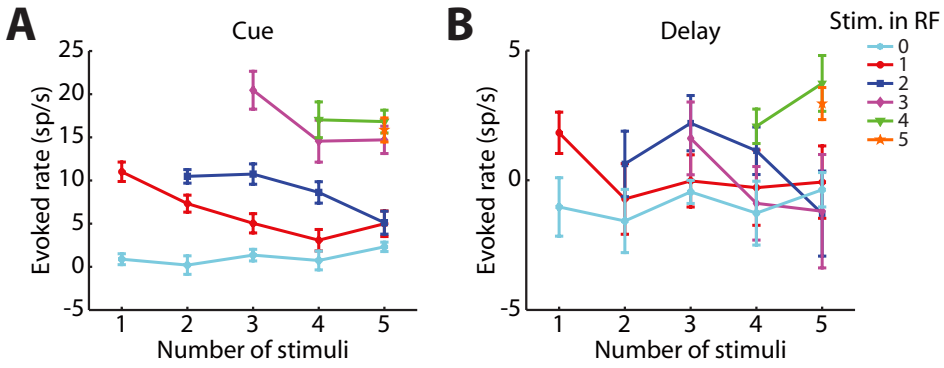

Fig. S4. Activity as a function of stimuli in and out of receptive field. Average firing rate evoked by displays with the same number of stimuli in receptive field, while different numbers of stimuli were presented out of the receptive field. Mean firing rate is shown for the cue period (A) and the delay period (B) after subtracting the firing rate averaged over the fixation period. Abscissa represents the total number of stimuli in each display. Different curves plot results from different numbers of stimuli appearing inside the respective field (shown in inset). Data from all neurons are included, whether they exhibited persistent activity in the delay period or not: n=415 displays with zero stimulus, n=448 for one stimulus, n=321 for two stimuli, n=184 for three stimuli, n=118 for four stimuli, n=33 for five stimuli. Evoked rates below zero in panel B (particularly the case when no stimulus appeared in the receptive field) represent mean activity in the delay period lower than the activity averaged during the entire fixation period, which often ramps in anticipation of the first stimulus and does not represent a true, stable "baseline" of the neuron's activity.

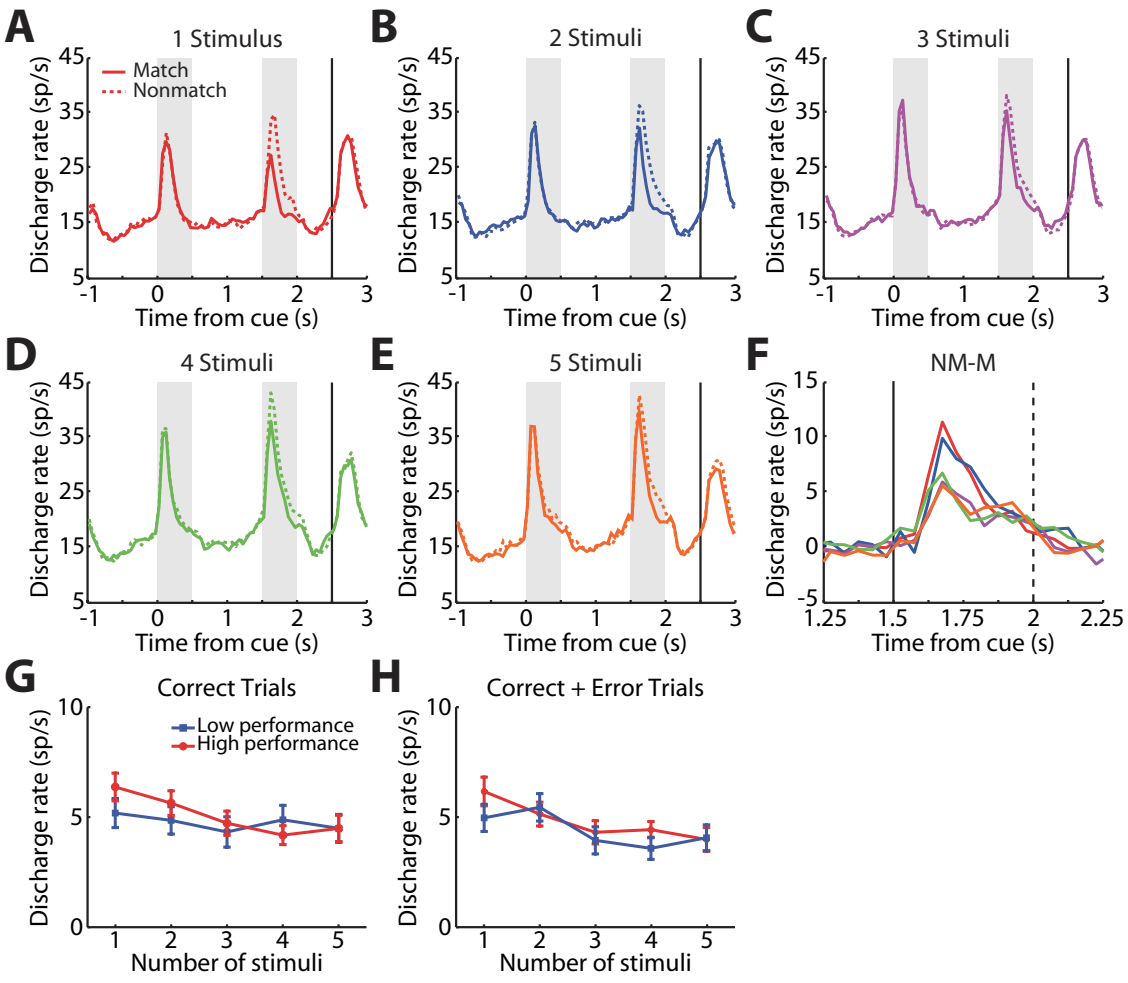

Fig. S5. Responses to Match and Nonmatch stimuli. (A-E) Mean firing rate in Match and Nonmatch trials for displays of varying stimuli. PFC neurons that were selective for the Match or Nonmatch status of the second stimulus presentation (2-way ANOVA, $p<0.01$) during the sample period were included in this analysis, n=127 (Monkey EL=92, Monkey DA=35). (F) Responses in the sample period are plotted for displays of different numbers of stimuli. (G) Averaged absolute value of the firing rate difference for Match and Nonmatch stimuli as a function of number of stimuli in the display during the sample period, for correct trials only. (H) As in G, for correct and error trials. Data from two monkeys, n=64 neurons with sufficient numbers of error trials for low performance, n=63 for high performance sessions.

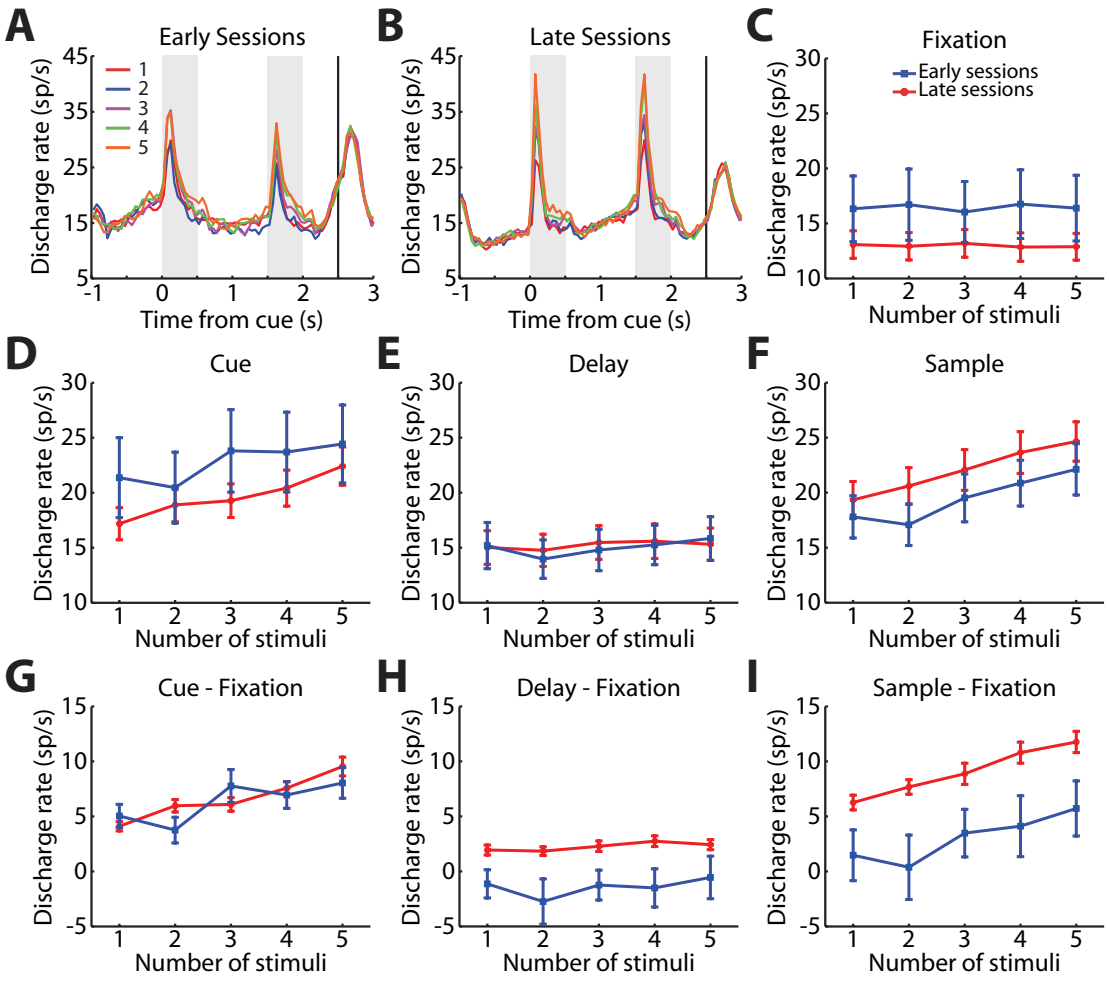

Fig. S6. Activity in early- and late- recording sessions. (A-B) Population PSTH obtained during presentation of multiple stimuli in early and late sessions based on a medial split. Data from two monkeys, n=111. (C-F) Averaged firing rate is shown for displays of different stimulus number for early and late sessions, during the fixation period (C), cue period (D), delay period (E) or sample period (F). (G-I) Average evoked firing rate of different stimuli number for early and late sessions during the cue period (G), delay period (H) or sample period (I).

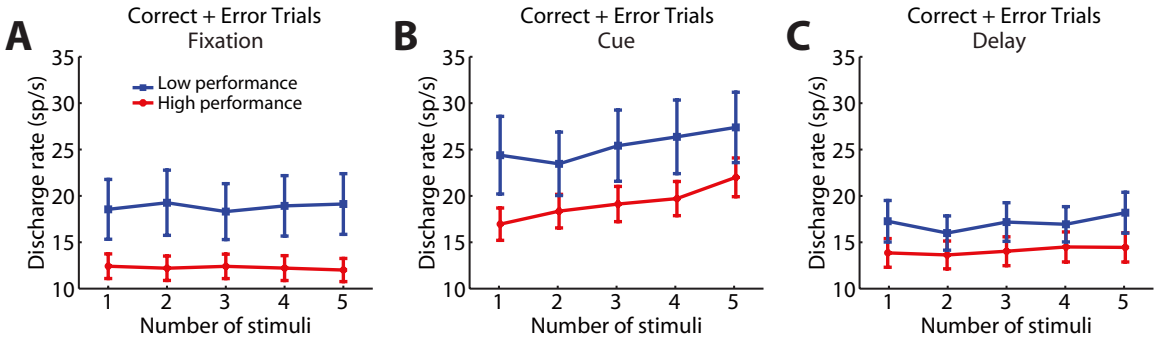

Fig. S7. Activity in low- and high-performance sessions including error trials. Averaged firing rate of different stimuli number for correct and error trials during the fixation period (A) cue period (B) or delay period (C). Two-way ANOVA, $p<0.01$. Data from two monkeys, n=111 neurons.

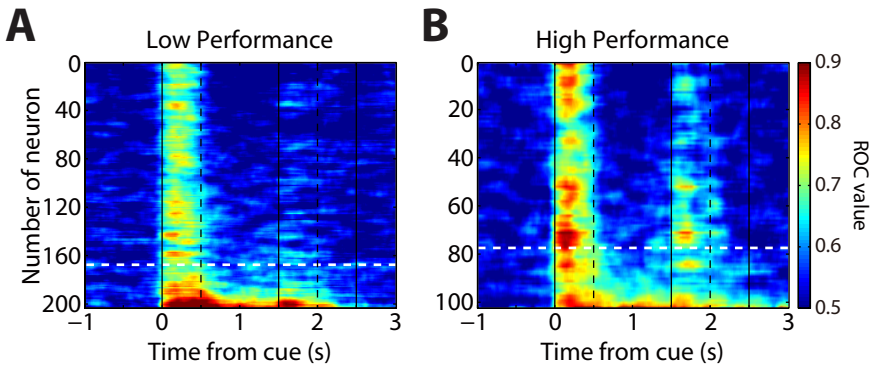

Fig. S8. ROC analysis for delay period. (A,B) ROC values for each recorded neuron in low- (A) and high-performance sessions (B) are shown, with sessions sorted based on higher ROC value achieved in the delay period. Dark red colors indicate high sensitivity between different stimulus displays of the same number. Dashed white lines indicate the ROC value of 0.6 averaged over the entire delay period. Data from two monkeys, n=305 neurons.

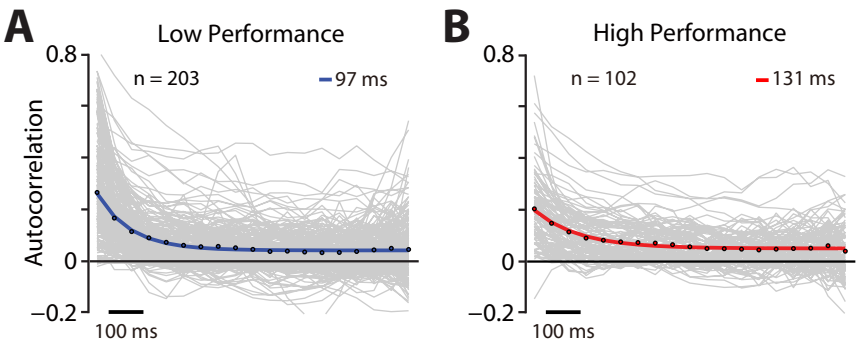

Fig. S9. Intrinsic time constant. (A) Autocorrelation of firing rate in fixation period for neurons recorded in low-performance sessions (n=203 neurons). Time constant of exponential fit is shown (τ=97 ms). (B) As in A, for neurons recorded in high-performance sessions (n=102 neurons). A significantly longer time constant (τ=131 ms) was now evident, (permutation test, p=7.77 x10-5).

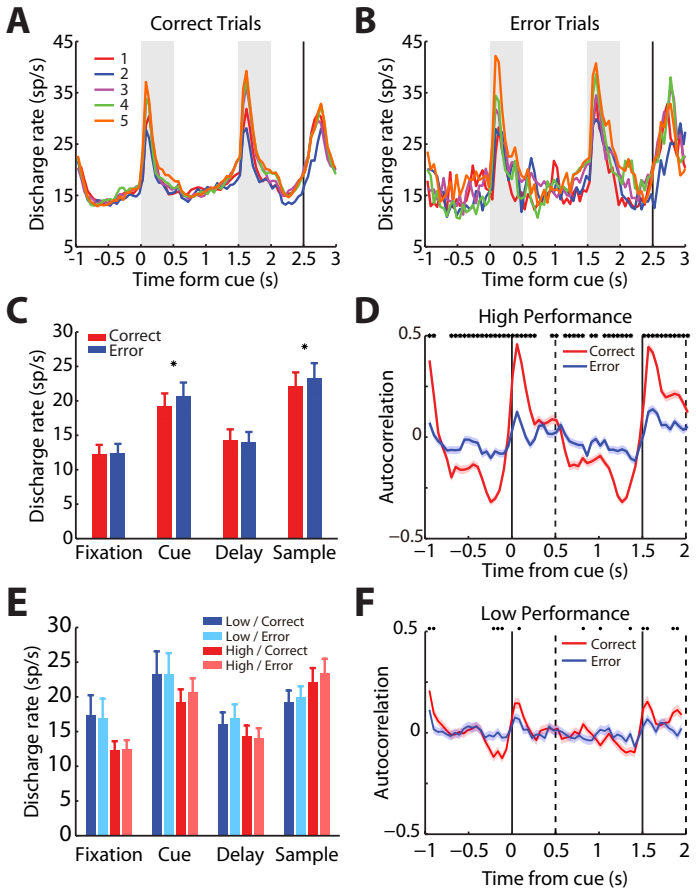

Fig. S10. Correct and Error trials. Population PSTHs depict activity in (A) correct and (B) error trials as a function of number of stimuli, obtained in high-performance sessions, (n=58 neurons). (C) Mean firing rate in each epoch in correct and error trials. Star represents significant difference at the p=0.05 level, based on paired t-test. (D). Autocorrelation of firing rate for correct and error trials. Correlation coefficient is plotted between firing rate in first 500 ms of fixation interval and successive 500 ms intervals, for correct and error trials of high-performance sessions (n=58 neurons). Asterisks at the top of the plot represent bins for which a significant difference ($p<0.05$) was present between correct and error trials. (E) Mean firing rate in each epoch in correct and error trials of high and low performance sessions. (F) Autocorrelation of firing rate for correct and error trials in low performance sessions (n=53 neurons).

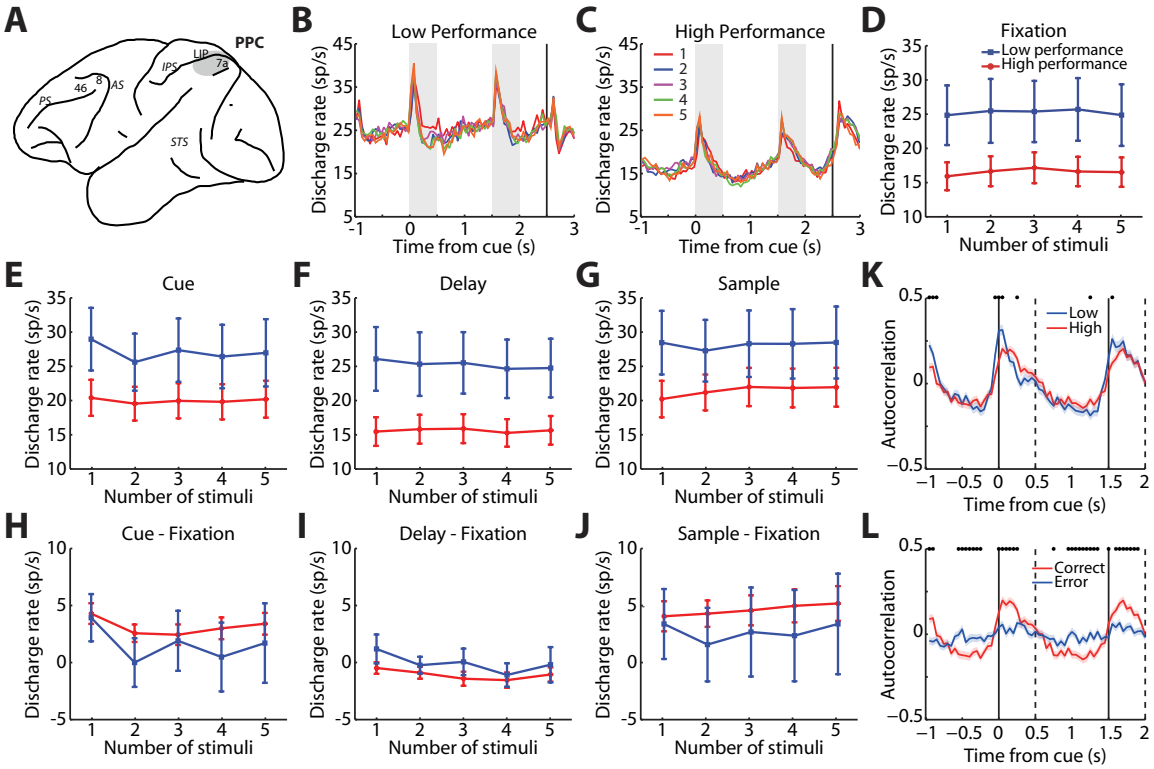

Fig. S11. Activity in Posterior Parietal Cortex. (A) Recordings in PPC sampled areas 7a and LIP. (B,C) Population PSTH obtained during presentation of multiple stimuli for high performance sessions (B) and low performance sessions (C). (D-G) Averaged firing rate of PPC neurons for different numbers of stimuli recorded in low- and high-performance sessions during the fixation period (D), cue period (E), delay period (F) or sample period (G). (H-J) Average evoked firing rate of different stimuli number for low- and high-performance sessions during the cue period (H), delay period (I) or sample period (J). 2-way ANOVA, $p<0.01$, n=39 neurons for high performance, n=33 for low performance. (K). Autocorrelation of firing rate for low- and high-performance sessions. (L) Autocorrelation of firing rate for correct and error trials. Only neurons that were selective to the stimuli (1-way ANOVA, $p<0.05$) during either the cue period or the delay period are included in this analysis. Data from monkey EL, n=72 neurons.

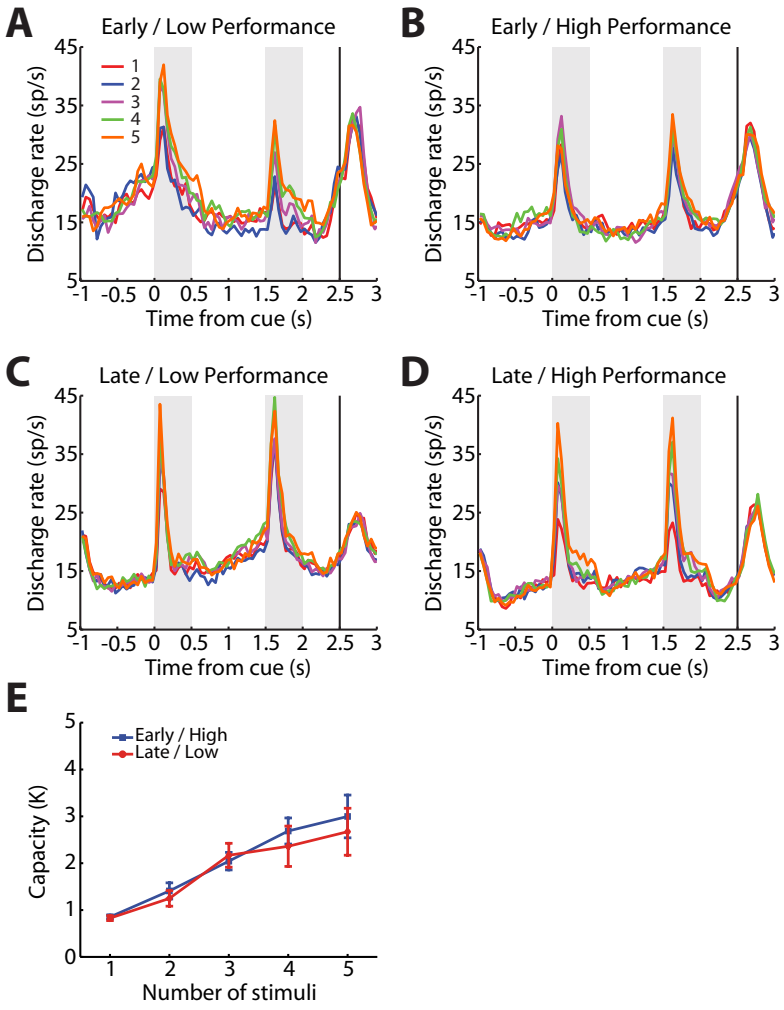

Fig. S12. Early and late sessions divided by low and high performance. Population PSTH of neurons recorded in (A) early sessions in which the monkeys achieved low performance (n=28); (B) early sessions in which the monkey achieved high performance (n=27); (C) late sessions in which the monkey achieved low performance (n=26); and (D) late sessions in which the monkey achieved high performance (n=30). (E) Working memory capacity in the sessions of panels B (early/high performance) and C (late/low performance) as a function of set size.

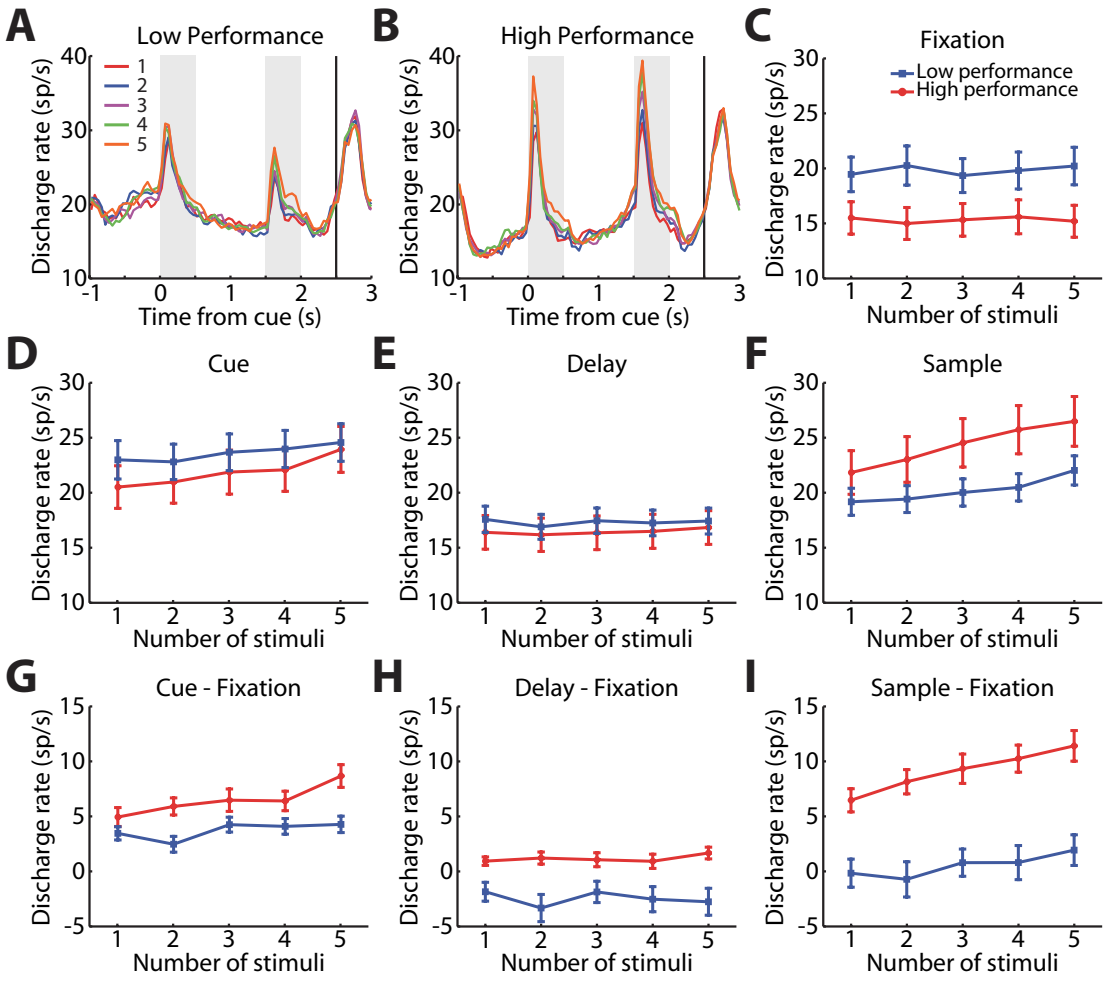

Fig. S13. Activity of all neurons in low- and high-performance sessions. (A-B) Population PSTH obtained during presentation of multiple stimuli in early and late sessions based on a medial split. Data from two monkeys, including all neurons, not only selective ones, n=305. (C-F) Averaged firing rate of different stimuli number for low- and high-performance sessions during the fixation period (C), cue period (D), delay period (E) or sample period (F). (G-I) Average evoked firing rate of different stimuli number for low- and high-performance sessions during the cue period (G), delay period (H) or sample period (I).

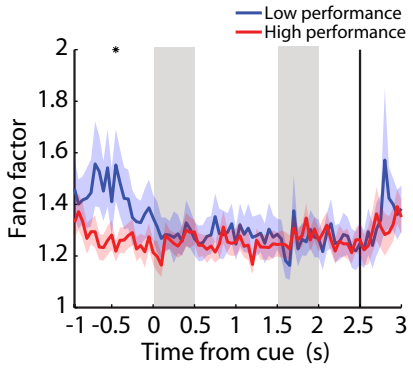

Fig. S14. Fano factor. Time-resolved plot of Fano factor (variance divided by mean) of spike counts in low- and high-performance sessions. The two curves are based on groups of neurons from the low and high-performance sessions that were matched for firing rate, in each bin. Shaded zone around each curve represents the standard error of the mean Fano factor value determined in this bin. Asterisks at the top of the plot represent bins for which a significant difference ($p<0.05$) was present between low and high-performance Fano factors.

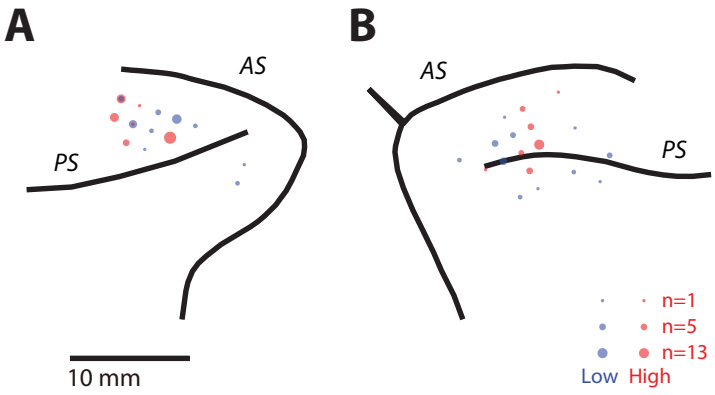

Fig. S15. Penetration maps. Anatomical location of where neurons were recorded in sessions of low performance (red) and sessions of high performance (blue). Size of circle represents number of neurons recorded at this track. Panels (A) and (B) depict data from the two monkeys, respectively. Horizontal bar = 10 mm. Abbreviations, AS: Arcuate Sulcus; PS: Principal Sulcus.

References

1. Meyer T, Qi XL, Stanford TR, & Constantinidis C (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J Neurosci* 31(17):6266-6276.
2. Meyer T & Constantinidis C (2005) A software solution for the control of visual behavioral experimentation. *J Neurosci Methods* 142(1):27-34.
3. Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10(4):433-436.
4. Tang H, Riley MR, & Constantinidis C (2017) Lateralization of Executive Function: Working Memory Advantage for Same Hemifield Stimuli in the Monkey. *Frontiers in neuroscience* 11:532.
5. Pashler H (1988) Familiarity and visual change detection. *Percept Psychophys* 44(4):369-378.
6. Harris KD, Henze DA, Csicsvari J, Hirase H, & Buzsaki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J Neurophysiol* 84(1):401-414.
7. Kobak D*, et al.* (2016) Demixed principal component analysis of neural population data. *eLife* 5.
8. Murray JD*, et al.* (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* 17(12):1661-1663.
9. Churchland MM*, et al.* (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13(3):369-378.
10. Qi XL & Constantinidis C (2012) Variability of prefrontal neuronal discharges before and after training in a working memory task. *PLoS ONE* 7(7):e41053.
11. Cavada C & Goldman-Rakic PS (1989) Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J. Comp. Neurol.* 287(4):422-445.
12. Qi XL & Constantinidis C (2015) Lower neuronal variability in the monkey dorsolateral prefrontal than posterior parietal cortex. *J Neurophysiol* 114(4):2194-2203.
13. Goris RL, Movshon JA, & Simoncelli EP (2014) Partitioning neuronal variability. *Nat Neurosci* 17(6):858-865.
14. Garavan H, Kelley D, Rosen A, Rao SM, & Stein EA (2000) Practice-related functional activation changes in a working memory task. *Microsc.Res.Tech.* 51(1):54-63.
15. Hempel A*, et al.* (2004) Plasticity of cortical activation related to working memory during training. *Am.J Psychiatry* 161(4):745-747.
16. Dahlin E, Neely AS, Larsson A, Backman L, & Nyberg L (2008) Transfer of learning after updating training mediated by the striatum. *Science* 320(5882):1510-1512.
17. Subramaniam K*, et al.* (2012) Computerized cognitive training restores neural activity within the reality monitoring network in schizophrenia. *Neuron* 73(4):842-853.
18. Jolles DD, van Buchem MA, Crone EA, & Rombouts SA (2013) Functional brain connectivity at rest changes after working memory training. *Hum Brain Mapp* 34(2):396-406.
19. Olesen PJ, Westerberg H, & Klingberg T (2004) Increased prefrontal and parietal activity after training of working memory. *Nat Neurosci* 7(1):75-79.
20. Qi XL, Meyer T, Stanford TR, & Constantinidis C (2011) Changes in Prefrontal Neuronal Activity after Learning to Perform a Spatial Working Memory Task. *Cereb. Cortex* 21:2722-2732.
21. Schneiders JA, Opitz B, Krick CM, & Mecklinger A (2011) Separating intra-modal and across-modal training effects in visual working memory: an fMRI investigation. *Cereb Cortex* 21(11):2555-2564.

22. Kuhn S, *et al.* (2013) The dynamics of change in striatal activity following updating training. *Hum Brain Mapp* 34(7):1530-1541.
23. Schweizer S, Grahn J, Hampshire A, Mobbs D, & Dalgleish T (2013) Training the emotional brain: improving affective control through emotional working memory training. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 33(12):5301-5311.
24. Takeuchi H, *et al.* (2013) Effects of working memory training on functional connectivity and cerebral blood flow during rest. *Cortex; a journal devoted to the study of the nervous system and behavior* 49(8):2106-2125.
25. Lippert MT, Steudel T, Ohl F, Logothetis NK, & Kayser C (2010) Coupling of neural activity and fMRI-BOLD in the motion area MT. *Magnetic resonance imaging* 28(8):1087-1094.
26. Wimmer K, Nykamp DQ, Constantinidis C, & Compte A (2014) Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat Neurosci* 17(3):431-439.
27. Standage D & Pare M (2018) Slot-like capacity and resource-like coding in a neural model of multiple-item working memory. *J Neurophysiol* 120(4):1945-1961.
28. Edin F, *et al.* (2009) Mechanism for top-down control of working memory capacity. *Proc Natl Acad Sci U S A* 106(16):6802-6807.
29. Schneegans S & Bays PM (2018) Drift in Neural Population Activity Causes Working Memory to Deteriorate Over Time. *J Neurosci* 38(21):4859-4869.
30. White JM, Sparks DL, & Stanford TR (1994) Saccades to remembered target locations: an analysis of systematic and variable errors. *Vision Res* 34(1):79-92.