



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

Hua Tang^{a,b,c}, Xue-Lian Qi^a, Mitchell R. Riley^{a,d}, and Christos Constantinidis^{a,1}

^aDepartment of Neurobiology & Anatomy, Wake Forest School of Medicine, Winston-Salem, NC 27157; ^bCenter for Neuropsychiatric Diseases, Institute of Life Science, Nanchang University, 330031 Nanchang, China; ^cNational Institutes of Mental Health, National Institutes of Health, Bethesda, MD 20892; ^dDepartment of Psychology, Vanderbilt University, Nashville, TN 37240

Edited by Ranulfo Romo, Universidad Nacional Autónoma de México, Mexico City, D.F., Mexico, and approved February 21, 2019 (received for review October 10, 2018)

The amount of information that can be stored in working memory is limited but may be improved with practice. The basis of improved efficiency at the level of neural activity is unknown. To investigate this question, we trained monkeys to perform a working memory task that required memory for multiple stimuli. Performance decreased as a function of number of stimuli to be remembered, but improved as the animals practiced the task. Neuronal recordings acquired during this training revealed two hitherto unknown mechanisms of working memory capacity improvement. First, more prefrontal neurons became active as working memory improved, but their baseline activity decreased. Second, improved working memory capacity was characterized by less variable temporal dynamics, resulting in a more consistent firing rate at each time point during the course of a trial. Our results reveal that improved performance of working memory tasks is achieved through more distributed activation and invariant neuronal dynamics.

working memory | monkey | executive function | neurophysiology | cognitive training

Working memory is the ability to maintain and manipulate information in mind (1). The capacity of human working memory is notoriously limited; only a handful of items can be held in memory over a period of seconds, creating a central limitation of human cognition (2). Individual abilities are reliant on working memory capacity (3, 4). Recent results suggest that capacity can be improved by training in working memory tasks (5, 6). The extent over which performance improvements generalize to tasks that were not part of the training has been a matter of intense debate (7, 8). No less contentious has been the idea that computerized training can improve cognitive function in healthy adults (9).

Human fMRI studies have produced conflicting results about the effects of training, with some studies suggesting increases (10–15) and others decreases in activity (16–19). The former are interpreted as reflecting a higher level of activation or recruitment of a larger cortical area, the latter as suggestive of improvements in efficiency (20). Humans are able to effectively reduce working memory load by grouping or “chunking” multiple stimuli (21). For example, a series of 10 digits comprising a phone number can be remembered more easily as a set of three groups of three to four numbers. The effects of training remain speculative, however, and the concept of efficiency is poorly defined at the neural level.

Working memory is thought to be mediated by persistent activity in a network of interconnected neurons behaving as a bump attractor and representing remembered stimuli in the peak of network activity, which may drift in time, resulting in loss of precision (22–24). Neurons that are excited by their preferred stimuli remain active in the delay period of a working memory task. When the capacity of the network is exceeded, information about an item may decay or merge with another item, resulting in loss of information about this item (*SI Appendix*, Fig. S1). The changes in neuronal activity that allow the network to increase its capacity are unknown. Computational studies simulating networks of neurons generating persistent activity show that improved capacity could be achieved through increased excitatory coupling, resulting in increased activity representing stimuli in the delay period or reduced external drive resulting in lower baseline

and stimulus-driven firing rate (25, 26). However, recent work has revealed considerable dynamics in the time course of delay-period activity, and their effects on models of capacity is unclear (27). Alternative mechanisms have also been proposed as the neural correlate of working memory, some of which do not depend on elevated activity during the delay period at all (28, 29). The site of information maintenance has also been debated, with some studies suggesting that information about stimuli is maintained in posterior areas, rather than the prefrontal cortex itself (30, 31). Our study sought to determine the changes in neural activity that effect improvement in working memory performance after practice.

Results

We trained two monkeys to perform a working memory task that required them to remember the spatial locations of multiple stimuli appearing on a visual display and to indicate if a second display with an equal number of stimuli was identical to the first (Fig. 1*A* and *B*). The monkeys’ performance declined monotonically as a function of the number of stimuli as the load of information maintained in working memory increased (Fig. 1*C* and *SI Appendix*, Fig. S2). The effect of number of stimuli on performance was highly significant (one-way ANOVA, $F_{4,545} = 35.5$, $P = 2.25 \times 10^{-26}$ for monkey EL; $F_{4,105} = 26.65$, $P = 2.90 \times 10^{-15}$ for monkey DA). Based on the pattern of correct and error responses, we were able to determine the working memory capacity K (Fig. 1*D*) in this task (defined as the set size multiplied by hits – false alarms and divided by 1 – false alarms). Behavioral sessions and neuronal recordings were collected over a period of several weeks, during which performance of the animals improved gradually (Fig. 1*E* and *F*). A linear regression of capacity on successive recording days showed a positive

Significance

Working memory capacity is notoriously limited to a handful of items, creating one of the central bottlenecks of human cognition, but can be improved by training. The neural basis of this improvement remains a matter of debate, as human imaging studies have produced contradictory results about the mechanisms that effect improved capacity. To resolve this controversy, we recorded neuronal activity from monkeys while they were being trained to improve their ability in maintaining multiple stimuli in memory. Our results reveal that improvement of working memory is effected by a more distributed activation of the prefrontal cortex and invariant temporal dynamics of neuronal activity. These changes render the prefrontal network more robust, allowing it to maintain more items in memory.

Author contributions: C.C. designed research; H.T., X.-L.Q., M.R.R., and C.C. performed research; H.T., X.-L.Q., and C.C. analyzed data; and H.T. and C.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence should be addressed. Email: cconstan@wakehealth.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1817278116/-DCSupplemental.

Published online March 15, 2019.

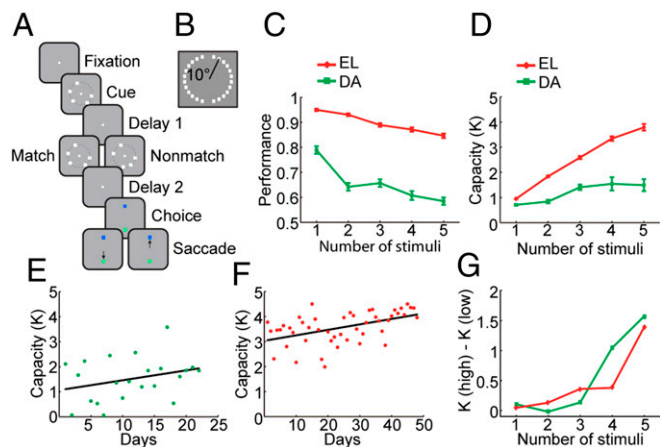


Fig. 1. Working memory capacity task and behavior. (A) Successive frames illustrate the sequence of events in the match/nonmatch task. The monkeys were required to remember the locations of all of the squares in the cue stimulus during a delay period. A second display then appeared, which contained the same number of stimuli. If one of the squares appeared at a new location, the display constituted a nonmatch; if the displays were identical, they constituted a match. Two choice targets of green and blue color appeared at the top and bottom location (randomly alternating in different trials), and the monkey was required to saccade to a green target if the two sequential displays matched each other or to a blue target otherwise. One animal was trained in a variation of this task that required a lever release for a matching stimulus instead of choice targets. (B) The 24 possible locations where stimuli could appear in the spatial match/nonmatch task. (C) The percentage of correct trials is shown as a function of the number of stimuli in the display for the two monkeys. (D) Estimated capacity in low- and high-performance sessions are plotted for the two monkeys. (E) Capacity estimated in successive daily sessions for monkey DA when neurophysiological recordings were also obtained. Line represents linear regression. (F) Capacity for monkey EL. (G) Difference in capacity between sessions of high and low performance, which were determined based on a median split separately for each monkey, plotted as a function of number of stimuli.

slope for each animal ($b = 0.022$, $F_{1,46} = 14.290$, $P = 4.5 \times 10^{-4}$ for monkey EL; $b = 0.040$, $F_{1,20} = 1.952$, $P = 0.178$ for monkey DA). We relied on a median split to distinguish between sessions of low and high performance based on estimated capacity. The monkeys achieved an overall performance level of 76.7% and capacity $K = 2.24$ items in the low-performance sessions and 82.6% correct trials and $K = 3.33$ in the high-performance sessions. Errors that differentiated sessions of high performance from those of low performance involved mostly displays with four to five stimuli, as evidenced by differences in capacity between these groups of sessions plotted as a function of number of stimuli (Fig. 1G).

Neural Responses to Multiple Stimuli. We recorded from areas 8 and 46 of the dorsolateral prefrontal cortex while the animals were performing the task (Fig. 2A). A total of 305 neurons were obtained (218 and 87 neurons from the 2 monkeys, respectively). Of those, 111 neurons ($n = 61$ for monkey EL, $n = 50$ for monkey DA) exhibited significant selectivity across different displays (one-way ANOVA, $P < 0.05$) and therefore could be informative about the displays that needed to be maintained in memory. We relied on these selective neurons for most analysis; results from all neurons are also reported in some figures and in the *SI Appendix*. The mean firing rate of selective neurons exhibited a highly dynamic time course, starting to increase before the first stimulus display even appeared in the screen (time -1 to 0 in Fig. 2B), peaking shortly after the appearance of the stimulus, decreasing further after the stimulus disappeared, but increasing again at the end of the delay period and during the second stimulus presentation (Fig. 2B and *SI Appendix*, Fig. S3).

Across the population of selective neurons, mean firing rate increased monotonically as a function of number of stimuli during the

cue presentation period (Fig. 2B and C). As the location of the stimuli was randomized in each session, displays with more stimuli were more likely to activate a neuron, and higher levels of activity were elicited across the population. The difference in firing rate between displays with different numbers of stimuli was highly significant (repeated-measures ANOVA, $F_{4,440} = 25.3$, $P = 7.0 \times 10^{-19}$). During the delay period (Fig. 2D), a generally higher firing rate was present for displays with more stimuli, but the effect was less consistent (repeated-measures ANOVA, $F_{4,440} = 3.88$, $P = 4.12 \times 10^{-3}$).

It is important to emphasize that this increase in firing rate for more stimuli appearing at randomized locations applied to the overall population activity pooled together. When we examined responses involving a single stimulus appearing in the receptive field of a neuron, activity decreased as additional stimuli were added outside the receptive field (*SI Appendix*, Fig. S4, red trace). The same was true when two stimuli appeared in the receptive field and additional stimuli appeared outside (*SI Appendix*, Fig. S4, dark blue trace), and so on. This is akin to effects of crowding and lateral inhibition (32, 33). The firing-rate difference reached statistical significance for the condition of one stimulus appearing in the receptive field as increasing numbers of stimuli were added outside (one-way ANOVA, $F_{4,443} = 5.38$, $P = 3.08 \times 10^{-4}$) and for two stimuli in the receptive field as increasing numbers of stimuli were added outside ($F_{3,317} = 3.41$, $P = 0.018$ for two stimuli; $F_{2,181} = 1.23$, $P = 0.293$ for three stimuli).

In addition to the cue and delay periods, we also examined neural activity during the presentation of the second display, and we distinguished between match and nonmatch responses (*SI Appendix*, Fig. S5). Across the population of prefrontal neurons, responses to the second presentation of stimuli were generally higher when the display constituted a nonmatch rather than a match, an effect often referred to as repetition suppression (34). This effect, too, was sensitive to the number of stimuli. The absolute difference between match and nonmatch responses decreased as a function of number of stimuli in the display (one-way ANOVA, $F_{4,630} = 2.73$, $P = 0.029$).

Firing Rate Changes Associated with Performance Improvement. As our neurophysiological recordings were obtained over a period

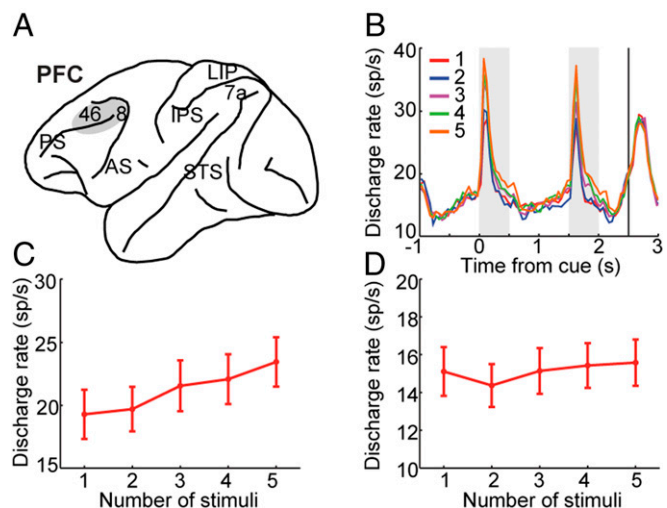


Fig. 2. Firing rate for displays of varying stimuli. (A) Schematic diagram of the monkey brain highlighting areas where recordings were performed. Recordings in dorsolateral prefrontal cortex (PFC) sampled areas 8 and 46. AS, arcuate sulcus; PS, principal sulcus. (B) PSTH represents mean population activity obtained during presentation of varying stimuli. (C and D) Mean firing rate is shown, averaged across displays of different numbers of stimuli, during the cue period (C) and the delay period (D). Data from two monkeys ($n = 111$ neurons that were selective to the stimuli during the cue period or the delay period). Bars represent SEM.

of time during which performance in the task improved (Fig. 1 *E* and *F*), it was possible to examine the neural changes that accompanied increased ability to maintain a greater memory load. A total of 203 neurons were recorded in low-performance sessions based on a median split depending on performance (as in Fig. 2*G*). Of those, 53 (26%) were selective for the stimulus pattern. A total of 102 neurons were recorded in high-performance sessions. The percentage of neurons selective for stimulus pattern increased substantially to 57% (58 of 102 neurons). This increase in percentage of selective neurons was statistically significant (χ^2 test, $P = 1.39 \times 10^{-7}$).

We hypothesized that, as performance of the working memory task improved, neuronal firing rate among neurons selective for the stimuli would also increase, as previous studies of the relationship between neuronal activity and working memory performance across sessions in simpler working memory tasks have shown (35). Unexpectedly, we found that prefrontal activation generally decreased in sessions of higher performance (Fig. 3 *A* and *B*). A two-way ANOVA with number of stimuli and low/high performance as factors revealed that the main effect of performance was significant in the fixation period ($F_{1,545} = 18.44$, $P = 2.08 \times 10^{-5}$; Fig. 3*C*), the cue period ($F_{1,545} = 10.85$, $P = 1.05 \times 10^{-3}$; Fig. 3*D*), and the delay period ($F_{1,545} = 5.86$, $P = 0.016$; Fig. 3*E*). For the sample period (Fig. 3*F*), prefrontal activation was slightly higher for sessions of higher performance, although the difference did not reach significance ($F_{1,545} = 3.27$, $P = 0.071$).

Similar results were obtained if we split sessions purely chronologically, rather than based on performance (SI Appendix, Fig. S6). These analyses were based on correct trials. As the improvement in performance involved a decrease in the rate of trials that end up being incorrect, we also repeated our analysis by using correct and error trials (SI Appendix, Fig. S7). The results including the error trials were essentially identical; a significant decrease in activity was present for the fixation period ($F_{1,545} = 18.73$, $P = 1.79 \times 10^{-5}$), the cue period ($F_{1,545} = 10.97$, $P = 9.90 \times 10^{-4}$), and the delay period ($F_{1,545} = 6.98$, $P = 8.48 \times 10^{-3}$) as a function of cumulative number of sessions.

The difference in activity between low- and high-performance sessions was already present from the baseline fixation period, we were therefore interested to examine the evoked neuronal firing rate relative to the baseline and to compare this measure between sessions of low and high performance (Fig. 3 *D–F*, dotted lines). Compared with low-performance sessions, no significant difference was found in sessions of higher performance in the cue period ($F_{1,545} = 0.37$, $P = 0.543$), but now a prefrontal activation increase was found in the delay period ($F_{1,545} = 22.49$, $P = 2.71 \times 10^{-6}$) and the sample period ($F_{1,545} = 49.45$, $P = 6.12 \times 10^{-12}$). The same effect was present if we divided sessions chronologically (SI Appendix, Fig. S6).

The combined effect of a greater proportion of neurons being activated and a lower level of absolute firing rate of these neurons resulted in a more distributed representation of stimulus information across the prefrontal population. This effect could be seen when we performed a receiver operating characteristic (ROC) analysis comparing responses for the best vs. worst display among displays of equal numbers of stimuli for all neurons recorded in the low- and high-performance sessions (Fig. 4 and SI Appendix, Fig. S8 for the delay period). A greater proportion of neurons achieved values of area under the ROC curve greater than 0.75 (a midpoint between chance and perfect performance) in the high-performance sessions.

We also examined the selectivity for match or nonmatch responses in low- and high-performance sessions (SI Appendix, Fig. S5*G*). Although, across the population, a substantial difference in firing rate was evident for match and nonmatch responses during the presentation of the second stimulus, this proved not to be predictive of low vs. high performance. A two-way ANOVA with factors number of stimuli and low/high performance group revealed no main effect of number of stimuli ($F_{4,625} = 1.81$, $P = 0.126$), performance ($F_{1,625} = 0.75$, $P = 0.386$), or interaction between number of stimuli and performance group ($F_{4,625} = 0.71$, $P = 0.584$). The results remained essentially unchanged when we repeated this analysis including error and correct trials ($F_{1,625} = 1.29$, $P = 0.256$ for main effect of performance; SI Appendix, Fig. S5*H*). A negligible difference was evident between high- and low-performance conditions particularly for five-stimulus displays, for which the greatest improvement in performance was observed. For this reason, and although the match/nonmatch effect was predictive of behavior across all sessions (SI Appendix, Fig. S5*F*), it could not account for the working memory performance improvement.

Changes in Neuronal Dynamics. Effects of working memory enhancement may not be limited only to mean firing rate changes between conditions. Indeed, the effects of dynamics in the representation of stimulus information have recently begun to be appreciated (27). We therefore proceeded to examine changes in the neural activity beyond simple changes in firing rate averaged across entire epochs. An additional difference between low- and high-performance sessions was that the envelope of firing rate changes became much more stereotypical in high-performance sessions [evidenced by the smoothness of curves in the population peristimulus time histogram (PSTH) of Fig. 3 *A* and *B*]. We quantified that change by performing a demixed principal component analysis (dPCA), which identifies components of firing rate related to different aspects of the task and stimuli (36). Similar to PCA, this technique reduces the dimensionality of the neuronal firing rate matrix, identifying the dimensions that capture the most variance in the neuronal population firing rate across conditions. Importantly,

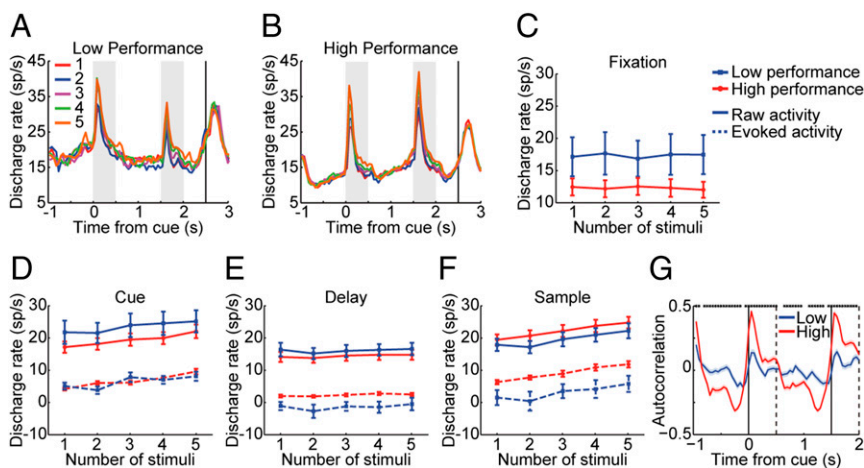


Fig. 3. Activity in low- and high-performance sessions. (A and B) Population PSTH obtained during presentation of multiple stimuli for (A) low-performance sessions and (B) high-performance sessions. Data from two monkeys ($n = 111$ neurons). (C) Averaged firing rate of different stimuli number for low- and high-performance sessions during the fixation period. (D–F) Raw firing rate and evoked firing rate after subtracting the baseline fixation rate in the cue period (D), delay period (E), and sample period (F). (G) Autocorrelation function plotting correlation coefficient between firing rate in the first 500 ms of fixation and every successive 500-ms interval for low- and high-performance sessions.

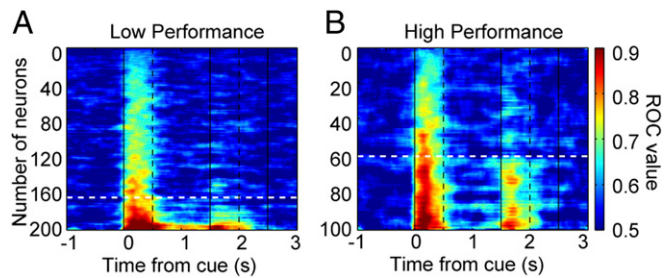


Fig. 4. ROC analysis. (A and B) ROC values for each recorded neuron in low-performance (A) and high-performance sessions (B) are shown. Dark red colors indicate high selectivity between different stimulus displays of the same number. Dashed white lines indicate the ROC value of 0.75. Data from two monkeys ($n = 305$ neurons).

dPCA segregates variance into components that take into account experimental parameters, which, in our case, included the time of stimulus presentation, the different displays used, and the categorical decision between match and nonmatch.

The results of this analysis indicated that “condition-independent” components (i.e., components unrelated to which stimulus display appeared on the screen or whether the trial involved a match or a nonmatch) represented 50% of the total firing rate variance in sessions of low performance and increased to 64% in sessions of high performance (pie charts in Fig. 5). In other words, a larger proportion of variance could be explained by changes in firing rate that tracked the time course of the trial (e.g., ramping in anticipation of the initial stimulus, cue transient, decline of activity in the delay period, sample transient). In contrast, components representing the stimulus and mixtures of stimulus and other parameters decreased from 46% in low-performance sessions to 32% in high-performance sessions (Fig. 5). The analysis also confirmed that representation of decision variables (essentially whether a stimulus was a match or nonmatch in the context of our task) changed little between low- and high-performance sessions: 24% of the total firing rate variance was accounted for by decision variables and mixtures in the low-performance, compared with 18% in the high-performance sessions.

As a result of the invariant dynamics of firing in high-performance sessions, the firing rate of a neuron during the fixation period was highly predictive of firing during the delay period in the high-performance sessions and did not simply reflect a

“baseline” level of activity. As a way to quantify the regularity of firing rate time course, we performed an autocorrelation analysis, measuring the correlation coefficient between the firing rate of a neuron at the first 500 ms of the fixation period and subsequent 500-ms intervals (Fig. 3G). Increases and decreases in firing rate were highly stereotypical, resulting in much greater positive and negative deviations in the high- vs. low-performance sessions. This effect could also be seen in the error bars of Fig. 3E (SI Appendix, Fig. S6H shows sessions split chronologically). The SD of delay period firing rate minus fixation firing rate was considerably lower in the high-performance ($\sigma = 3.2$ spikes per second) than in the low-performance sessions ($\sigma = 12.7$ spikes per second). Note that variability of the absolute rate of firing during the delay period was not appreciably greater in the high- and low-performance sessions (error bars for solid lines vs. dotted lines in Fig. 3E).

As the differences in firing patterns between low- and high-performance sessions were already evident in the fixation period, before the appearance of the first stimulus display, we calculated the time constant of the autocorrelation of firing rate in the fixation period (SI Appendix, Fig. S9). This is a measure of the neuron’s intrinsic time scale that quantifies the stability of firing dynamics (37). The time constant increased from 97 ms in the low-performance sessions to 131 ms in the high-performance ones, a significant difference (permutation test, $P = 7.77 \times 10^{-5}$).

Relationship with Behavior. Not all changes we identified between low- and high-performance sessions were predictive of behavior in the task. To identify the critical aspects of neuronal activity that were associated with performance, we compared activity patterns in correct and error trials in the high-performance sessions (SI Appendix, Fig. S10A and B). Firing rate differences were generally subtle between correct and error trials. Whereas a huge decrease in baseline activity was evident in high-performance compared with low-performance sessions (Fig. 3), the average rates in correct and error trials were very similar (SI Appendix, Fig. S10C) and no significant difference was evident (paired t test, $t_{58} = 0.4563$, $P = 0.650$). Collapsed across all stimulus conditions, trials when a cue elicited higher activity were slightly more likely to result in errors (paired t test, $t_{58} = 2.835$, $P = 6.33 \times 10^{-3}$). The same was true for the sample presentations (SI Appendix, Fig. S10C). The strongest predictor of performance was the regularity of the firing rate time course. The correlation coefficient between firing rate during the fixation period and every other time point during the trial was greatly reduced (i.e., flattened) in error trials (SI Appendix, Fig. S10D).

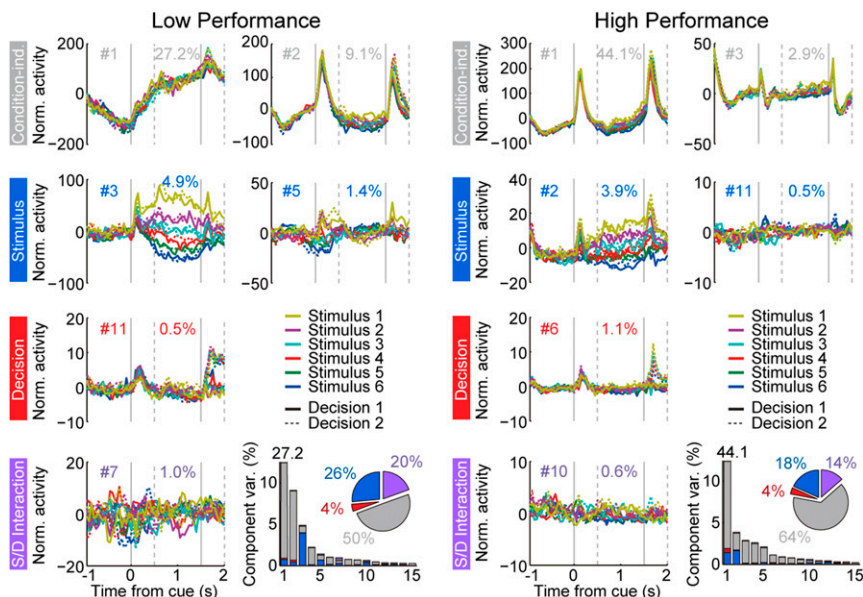


Fig. 5. dPCA analysis of low-performance (Left) and high-performance sessions (Right). The top row of graphs represents the first two condition-independent components of dPCA analysis, the second row represents the first two stimulus-related dPCA components, the third row represents the first decision-related dPCA component, and the bottom row represents the first stimulus/decision mixture dPCA component. Data from two monkeys ($n = 199$ neurons with sufficient numbers of trials for this analysis from low-performance sessions; $n = 92$ for high-performance sessions).

In other words, trials in which firing rate did not ramp up or down during the time course of the trial were more likely to result in errors. The differences between sessions of high vs. low performance and correct vs. error trials were additive; the most pronounced dynamics were observed in correct trials of high-performance sessions and the least in error trials of low-performance sessions (*SI Appendix, Fig. S10 D–F*).

Regional Specificity of Efficiency-Related Changes. A decrease in activity in the prefrontal cortex after training may imply that information about the location of stimuli is represented in activity of more posterior areas, as some current theories of working memory suggest (30, 38). To test this idea, we examined the activity of posterior parietal cortex (area 7a and lateral intraparietal area), a brain region providing direct afferent input to the prefrontal cortex (39). Activity changes in the posterior parietal cortex mirrored those in the prefrontal cortex. Baseline activity in the posterior parietal cortex also decreased in sessions of high performance vs. lower performance, and invariant temporal dynamics were also predictive of correct and error trials (*SI Appendix, Fig. S11*).

Discussion

Our study examined the neural basis of improvement in a task that required monkeys to maintain multiple stimuli in working memory. Improvement in performance was not perfectly monotonic with respect to practice. We focused primarily on sessions in which the monkeys had accumulated training and achieved the benefits of it; however, essentially identical conclusions were drawn when we split daily sessions based on level of performance reached or cumulative amount of practice. Three types of changes were associated with enhanced working memory performance. First, a greater percentage of neurons exhibited selectivity for the stimuli. Second, selective neurons were characterized by decreased firing rate. This was evident from the baseline, fixation interval of the task. However, no decrease in firing rate relative to baseline was evident for the delay and sample periods. The combined effect of the increase in the number of selective neurons and lower average firing rate resulted in a more distributed representation of information about the stimuli across the population. The decrease in firing rate was not exclusive to the prefrontal cortex; posterior parietal neurons exhibited similar changes. Third, improved performance was characterized by a more stereotypical time course of neuronal firing, with a greater proportion of variance being accounted for by condition-independent components. This effect was predictive of performance after training and appeared to be more pronounced in the prefrontal cortex. Our results suggest that improved efficiency of working memory representations rely on a more distributed representation of information across the network of neurons, silencing of background levels of activity, and more invariant temporal activation.

Effects of Working Memory Practice. Evidence exists that training can improve working memory and that performance increases emerge not only for the trained tasks, but also for tasks that were not part of the training; that is, there is transfer from one task to another (5, 6, 40). Some of these findings have been challenged, and the benefits of working memory training for healthy adults remain unclear (9). Nonetheless, recent meta-analyses have estimated at least a modest benefit of working memory improvement (7, 8). Furthermore, there is strong evidence that working memory training is beneficial for clinical populations, including children with attention deficit/hyperactivity disorder, stroke patients, and those with schizophrenia (5, 13, 41).

Training may improve efficiency with respect to management of working memory content and capacity for number of items maintained in memory (20). Strategies such as chunking allow subjects to group multiple stimuli in fewer items that can be more efficiently stored and maintained in memory (21). In the case of multistimulus visual displays, subjects may group

identical items rather than attempt to recall all of them independently (42, 43). Factors such as grouping, Gestalt principles of proximity and connectedness, and the specific arrangement of stimuli are known to confer an advantage for groups of stimuli maintained in memory over stimuli distributed between fields (44, 45). Even perception of visual scenes that are physically present relies on extraction of summary statistics that simplify the informational content of complex scenes (46). In the context of the task we trained monkeys to perform, efficiencies may be achieved by mentally transforming the display of multiple squares into a mental “polygon” rather than maintaining in memory every item separately (47). Our results demonstrate that such a more efficient representation of information in memory has tangible effects on neural activity. Firing rate during the trial decreased to a large extent. Changes in activity were not limited to the stimulus encoding during the stimulus presentation period, but preceded the stimulus and followed it, in the period of working memory maintenance. Human imaging studies have produced contradicting results on the effects of working memory training (20), with some studies suggesting increases (10–15) and others decreases in activity (16–19). Our results provide a more detailed picture at the neural population level and offer insights in these results. We document that more prefrontal neurons were responsive after practice, with a lower overall level of baseline activity in sessions during which performance improved.

Neural Basis of Working Memory and Its Capacity Limitation. In addition to efficiency, working memory capacity can be improved with training. Working memory is thought to be mediated by persistent activity generated during the delay interval of working memory tasks, but this issue, too, has been a matter of debate (28–30, 38). A well-known effect not dependent on persistent activity is differential responses to stimuli that match or do not match a previously presented one (34). Those match/nonmatch-differentiating responses have been shown to be predictive of behavior in previous studies (48), and we documented the same effect here (*SI Appendix, Fig. S5*). However, we saw no systematic changes in the ability of neurons to discriminate between match and nonmatch stimuli based on level of performance (*SI Appendix, Fig. S5 G and H*), and no greater percentage of variance across the population could be attributed to decision variables in high-performance sessions (Fig. 5). Therefore, working memory capacity improvement could not be mediated by such activity-silent mechanisms.

On the contrary, persistent activity generated in the prefrontal cortex represents information about stimuli maintained in memory, is predictive of overall performance in the task, and can account for the behavioral output of recall (23). The persistent-activity model of spatial working memory posits that appearance of a stimulus generates activity that is maintained during the delay period, but may drift from the original location over time (24). The location recalled by the subject is precisely determined by the drift of the delay period activity (*SI Appendix, Fig. S1*). The activity for each of multiple stored items held in memory can also be thought of as a bump attractor (26). As more items are added to memory, the total population activity increases until the capacity of the system is exceeded, the persistent activity representing some stimuli decays, and the respective stimuli cannot be recalled at the end of the delay period (26). Previous neurophysiological studies examining activity in working memory tasks involving multiple stimuli have shown that activity generally decreases as additional stimuli are added in the visual field as a result of factors such as lateral inhibition, and information about stimuli represented in delay-period activity quickly saturates (31, 49, 50). Our present results are consistent with this interpretation. Here we show that stimulus-selective persistent activity is generated in a multistimulus working memory task. After practice that leads to performance improvement, the absolute level of delay-period activity decreases relative to the earlier stage; however, persistent activity remains elevated relative to the baseline. This is beneficial for the stability of a network that

requires maintenance of multiple peaks of activity in different neuronal populations, as predicted by computational studies showing that reduced external drive that results in lower baseline and stimulus-driven firing rate can improve the robustness of the network (26). Such task-specific transformation of neuronal activity can also account for psychophysical advantages such related to grouping and chunking of multiple stimuli (44, 45). Somewhat analogous decreases in activity have also been reported in the context of dual task performance, requiring the maintenance in working memory not simply of multiple stimuli, but of multiple tasks (51).

A final factor affected by training was the temporal dynamics of neuronal firing. It has been recently recognized that prefrontal neurons exhibit strong temporal dynamics, which, however, coexist with stable population coding of the remembered stimulus, as trajectories of neuronal population firing remain separated in parameter subspace (27). Our results suggest that reducing the

variability of neuronal dynamics enhances the stability of stimulus representation in this coding scheme.

Experimental Procedures

Two male rhesus monkeys (*Macaca mulatta*) weighing 7–13 kg were used in this study. The monkeys were previously trained to perform working memory tasks with a single stimulus (35, 52). All experimental procedures followed guidelines by the US Public Health Service Policy on Humane Care and Use of Laboratory Animals and the National Research Council's Guide for the Care and Use of Laboratory Animals and were reviewed and approved by the Wake Forest University Institutional Animal Care and Use Committee. A detailed description of the study methodology is provided in *SI Appendix, Experimental Procedures*.

ACKNOWLEDGMENTS. We thank B. Averbeck, D. Blake, A. Compte, T. Klingberg, and T. Stanford for valuable comments. Research reported in this paper was supported by the National Eye Institute of the National Institutes of Health under Awards R01 EY016773, R01 EY017077, and R01 MH116675.

- Baddeley A (2012) Working memory: Theories, models, and controversies. *Annu Rev Psychol* 63:1–29.
- Cowan N (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci* 24:87–114, discussion 114–185.
- Kane MJ, et al. (2007) For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychol Sci* 18:614–621.
- Gathercole SE, Brown LH, Pickering SJ (2003) Working memory assessments at school entry as longitudinal predictors of national curriculum attainment levels. *Educ Psychol* 70:177–194.
- Klingberg T, Forssberg H, Westerberg H (2002) Training of working memory in children with ADHD. *J Clin Exp Neuropsychol* 24:781–791.
- Jaeggi SM, Buschkuhl M, Jonides J, Perrig WJ (2008) Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci USA* 105:6829–6833.
- Cortese S, et al.; European ADHD Guidelines Group (EAGG) (2015) Cognitive training for attention-deficit/hyperactivity disorder: Meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J Am Acad Child Adolesc Psychiatry* 54:164–174.
- Schwaighofer M, Fischer F, Buhner M (2015) Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educ Psychol* 50:138–166.
- Owen AM, et al. (2010) Putting brain training to the test. *Nature* 465:775–778.
- Garavan H, Kelley D, Rosen A, Rao SM, Stein EA (2000) Practice-related functional activation changes in a working memory task. *Microsc Res Tech* 51:54–63.
- Hempel A, et al. (2004) Plasticity of cortical activation related to working memory during training. *Am J Psychiatry* 161:745–747.
- Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L (2008) Transfer of learning after updating training mediated by the striatum. *Science* 320:1510–1512.
- Subramaniam K, et al. (2012) Computerized cognitive training restores neural activity within the reality monitoring network in schizophrenia. *Neuron* 73:842–853.
- Jolles DD, van Buchem MA, Crone EA, Rombouts SA (2013) Functional brain connectivity at rest changes after working memory training. *Hum Brain Mapp* 34:396–406.
- Olesen PJ, Westerberg H, Klingberg T (2004) Increased prefrontal and parietal activity after training of working memory. *Nat Neurosci* 7:75–79.
- 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* 21:2555–2564.
- Kühn S, et al. (2013) The dynamics of change in striatal activity following updating training. *Hum Brain Mapp* 34:1530–1541.
- Schweizer S, Grahn J, Hampshire A, Mobbs D, Dalgleish T (2013) Training the emotional brain: Improving affective control through emotional working memory training. *J Neurosci* 33:5301–5311.
- Takeuchi H, et al. (2013) Effects of working memory training on functional connectivity and cerebral blood flow during rest. *Cortex* 49:2106–2125.
- Constantinidis C, Klingberg T (2016) The neuroscience of working memory capacity and training. *Nat Rev Neurosci* 17:438–449.
- von Bastian RC, Oberauer K (2014) Effects and mechanisms of working memory training: A review. *Psychol Res* 78:803–820.
- Constantinidis C, et al. (2018) Persistent spiking activity underlies working memory. *J Neurosci* 38:7020–7028.
- Riley MR, Constantinidis C (2016) Role of prefrontal persistent activity in working memory. *Front Syst Neurosci* 9:181.
- Wimmer K, Nykamp DQ, Constantinidis C, Compte A (2014) Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat Neurosci* 17:431–439.
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10:910–923.
- Edin F, et al. (2009) Mechanism for top-down control of working memory capacity. *Proc Natl Acad Sci USA* 106:6802–6807.
- Murray JD, et al. (2017) Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc Natl Acad Sci USA* 114:394–399.
- Lundqvist M, Herman P, Miller EK (2018) Working memory: Delay activity, yes! Persistent activity? Maybe not. *J Neurosci* 38:7013–7019.
- Stokes MG (2015) 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends Cogn Sci* 19:394–405.
- D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. *Annu Rev Psychol* 66:115–142.
- Lara AH, Wallis JD (2014) Executive control processes underlying multi-item working memory. *Nat Neurosci* 17:876–883.
- Angelucci A, Levitt JB, Lund JS (2002) Anatomical origins of the visual receptive field and modulatory surround field of single neurons in macaque visual cortical area V1. *Prog Brain Res* 136:373–388.
- Pelli DG (2008) Crowding: A cortical constraint on object recognition. *Curr Opin Neurobiol* 18:445–451.
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23.
- 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.
- Kobak D, et al. (2016) Demixed principal component analysis of neural population data. *eLife* 5:e10989.
- Murray JD, et al. (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* 17:1661–1663.
- Sreenivasan KK, Curtis CE, D'Esposito M (2014) Revisiting the role of persistent neural activity during working memory. *Trends Cogn Sci* 18:82–89.
- 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* 287:422–445.
- Klingberg T, et al. (2005) Computerized training of working memory in children with ADHD—A randomized, controlled trial. *J Am Acad Child Adolesc Psychiatry* 44:177–186.
- Westerberg H, et al. (2007) Computerized working memory training after stroke—A pilot study. *Brain Inj* 21:21–29.
- Peterson DJ, Gözenman F, Arciniegua H, Berryhill ME (2015) Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Atten Percept Psychophys* 77:2270–2283.
- Xu Y, Chun MM (2007) Visual grouping in human parietal cortex. *Proc Natl Acad Sci USA* 104:18766–18771.
- Peterson DJ, Berryhill ME (2013) The Gestalt principle of similarity benefits visual working memory. *Psychon Bull Rev* 20:1282–1289.
- Xu Y (2006) Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Percept Psychophys* 68:815–828.
- Whitney D, Yamanashi Leib A (2018) Ensemble perception. *Annu Rev Psychol* 69:105–129.
- Tang H, Riley MR, Constantinidis C (2017) Lateralization of executive function: Working memory advantage for same hemifield stimuli in the monkey. *Front Neurosci* 11:532.
- Qi XL, Meyer T, Stanford TR, Constantinidis C (2012) Neural correlates of a decision variable before learning to perform a match/non-match task. *J Neurosci* 32:6161–6169.
- Buschman TJ, Siegel M, Roy JE, Miller EK (2011) Neural substrates of cognitive capacity limitations. *Proc Natl Acad Sci USA* 108:11252–11255.
- Matsushima A, Tanaka M (2014) Different neuronal computations of spatial working memory for multiple locations within versus across visual hemifields. *J Neurosci* 34:5621–5626.
- Watanabe K, Funahashi S (2014) Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nat Neurosci* 17:601–611.
- Katsuki F, Constantinidis C (2012) Early involvement of prefrontal cortex in visual bottom-up attention. *Nat Neurosci* 15:1160–1166.