

Supplementary Table 1

region	N (neurons)	time (hours)			waveform SNR			firing rate (Hz)		
LPFC	359	3.2	3.3	3.6	5.3	7.7	11.1	0.7	2.2	5.7
FEF	105	3.2	3.5	3.7	5.2	8.0	11.4	1.0	2.5	5.9
Parietal (7a/b)	193	3.2	3.5	3.8	5.5	7.6	10.4	0.8	2.0	5.1
V4	224	3.3	3.5	3.9	6.2	8.4	11.3	0.8	2.2	6.0

Supplementary Table 1. Count, recording time, waveform signal-to-noise ratio, and cue-evoked firing rate (post-cue epoch: 0-500 ms, averaged across all conditions) for the population of neurons from each region entering the classification analysis (Figure 2). Triplets reflect 25th/50th/75th percentiles across the population.

Supplementary Discussion 1: Balancing between generalized and task-specific representations for the control of attention and selection

In the main text, we show that a classifier trained to decode the location of selection from the activity of neurons in LPFC was able to generalize and decode the location of attention (and vice-versa; Fig. 2c, purple traces). These generalized representations may allow for behaviors learned in one domain (e.g., adding visible numbers) to be easily extended to a different domain (e.g., adding remembered numbers)¹.

However, it is important to note that we also found task-specific representations. In LPFC, the performance of the generalized classifier was not exactly the same as the selection classifier (Fig. 2c, blue traces; trained and tested within the selection condition alone). Similarly, we observed less generalization in FEF and no generalization in parietal or V4 (Fig. 2c). Previous work has suggested task-specific representations could arise from 'non-linear mixed selectivity'². For example, in the current task, neurons with non-linear mixed selectivity would respond uniquely to the combination of the cued location (upper vs. lower) and task (selection vs. attention), which would lead to a classifier being unable to generalize across tasks. While such non-linear, task-specific representations may hinder generalization, they have the advantage of increasing the coding capacity of the network². This allows associations to be learned for a specific stimulus, in a specific context. In contrast, generalized representations may emerge from linear mixed selective representations³, which have a lower coding capacity but allow learning to be generalized across contexts. Our results suggest the brain uses a combination of both generalized and specialized representations for selection and attention. This may allow the brain to balance the relative advantages (and disadvantages) of the two types of representations, generalizing learning when beneficial while also using task-specific representations to optimize behavior in a specific task.

Finally, we want to note that our results are correlative, and future work is needed to causally establish the role of prefrontal cortex in selection and attention. Future work is also needed to test whether generalized task representations are seen when selecting stimuli/memories in other tasks.

Supplementary Discussion 2: Memory information was distributed across prefrontal, parietal, and visual cortex

As described in the main text, individual neurons in LPFC, FEF, parietal and V4 carried information about the color of the upper and lower stimuli (Extended Data Fig. 5a). On average across the population, all four regions carried significant information about the color of the stimuli during the stimulus presentation (Fig. 3, left panels). Color information was then maintained across this distributed network during the first memory delay (Fig. 3, middle panels, before selection).

Previous work has argued for little working memory information in prefrontal cortex⁴ and visual cortex⁵ of monkeys (but see⁶). The increased information observed in our task could reflect a difference in how memories are stored; recent theoretical work⁷ suggests active representations are particularly important when manipulating memories, such as in the retro task, while activity-silent mechanisms may suffice for more passive tasks. Consistent with this, there was less information about the color of the remembered stimulus on pro trials in comparison to retro trials (Extended Data Fig. 6).

Supplementary Discussion 3: Estimating and comparing the transformation of the selected and unselected items

As described in the main text, selection transformed the representation of the selected item into a new subspace where the neural representations of the lower and upper items were aligned. This could facilitate read-out from a common subspace within LPFC in the post-cue period, allowing color to be decoded, regardless of the location of the item. In addition, the neural representation of the unselected item was also transformed, such that, in the post-cue period, it shifted into a new ‘unselected’ subspace that was orthogonal to the ‘selected’ subspace (Extended Data Fig. 9).

Given that both items were transformed, we were interested in understanding how the transformation of the selected item related to the unselected item. One hypothesis is that selection is a ‘united’ process that applies a single affine transformation to both the selected and unselected items, shifting them from their pre-cue representation to post-cue representation. Alternatively, the selection process could transform the selected and unselected memories independently, by applying a unique transform to each item (e.g., serially over time) or by applying a non-affine transformation. Finally, selection may involve a combination of both a united transform and independent transformations.

To understand if selection transformed the selected and unselected item in the same way, we estimated the transformation matrix that mapped a pre-cue representation onto its post-cue state (see methods). The transformation within a condition was stable: applying the estimated transformation to withheld pre-cue data from the same condition, accurately predicted the post-cue response (measured as a low reconstruction error between the predicted and observed post-cue representations, ED Fig. 9e, first column).

To discriminate the unified and independent hypotheses, we applied the estimated transform from the selected item (e.g., selected upper) to held-out data from the associated unselected item in the same trial condition (e.g., non-selected lower). As seen in Extended Data Figure 9e (second column), the transformation of the selected and unselected items was more similar than expected by chance (as indicated by a lower reconstruction error than in randomly permuted data). This suggests there is a common component that is transforming both the selected and unselected items in memory. However, this reconstruction error is also greater than what is seen within a condition (first column), suggesting that the transformation may also contain item-specific components (or may not be affine).

An alternative explanation for these results is that the united part of the transformation reflects the shift of representations from a general ‘pre-cue’ subspace (with both upper and lower items) to a general ‘post-cue’ subspace (with both the selected and unselected items). If this were true, then all transformations from pre- to post-cue should have low reconstruction error. To test this, we measured whether the transformation of an item at a specific location was the same regardless of whether it was selected or unselected (e.g., selected upper vs. non-selected upper). As seen in Extended Data Figure 9e (third column), the reconstruction error was not significantly different than chance (if anything, it was slightly greater than chance). Similarly, the transformation of an item did not generalize across both selection-condition and item (e.g., selected upper vs. selected lower or non-selected upper vs. non-selected lower). Again, the reconstruction error was not significantly different than chance (ED Fig. 9e, fourth column). Together these results suggest the common component of the transformation was not simply due to a generic change in memory representations from a pre-cue subspace to a post-cue subspace. Rather, selection seems to act, in part, on both the selected and unselected memory

representations, transforming them from independent, item-specific, subspaces to an aligned 'template' subspace, facilitating visual search and the behavioral report.

Supplementary Discussion 4: Cognitive control through dynamic transformations

As detailed in the main text, and as previously theorized⁸, we found selection transformed memory information in a manner consistent with task demands. Early in the trial, working memory representations were held in item-specific ‘upper’ and ‘lower’ spaces, likely to facilitate the selection of a memory by its associated location. Then, later in the trial, the selected memory shifted into a shared ‘template’ space, which could be used to guide responses by acting as a template for searching the color wheel. Previous work has found stable subspaces can maintain working memory information despite dynamics in neural activity⁹; our work suggests multiple such subspaces can exist and that information can be transformed between subspaces in a task-dependent manner⁸.

Importantly, all three spaces (upper, lower, and template) were approximately orthogonal to one another, which could reduce interference between simultaneously maintained memory representations (i.e. upper and lower) and limit interactions between memory representations and search-related representations.

The dynamic transformation of the selected memory from the upper/lower space to the shared template space is reminiscent of the rotation of representations from a passive ‘null’ space to an active ‘response’ space in motor planning and attention^{10,11}. Our results are consistent with this work and extend it to show multiple representational spaces can converge onto a single common space (i.e., both lower and upper transform into the template space). Furthermore, we find these dynamics are under cognitive control and depend on task demands, reflected in the fact that the transformation happens after selection during the retro task but immediately in the pro (attention) task.

More broadly, dynamic transformations could be a mechanism of cognitive control. Cognitive control is thought to rely on task-specific routing of information, either due to gain modulation¹² or changes in synchrony^{13,14}. Our results suggest an additional mechanism – cognitive control dynamically transforms information in a task-specific manner, allowing information to selectively engage with task-relevant circuits¹⁵. For example, in our task, a downstream ‘visual search’ circuit could use color information from the common template space to guide visual search. Early in the trial, this circuit is not engaged, as memories are stored in the orthogonal upper/lower spaces and so colors are not differentiable by the visual search circuit (Fig. 4g, left; i.e., they are in the circuits null space). Later, selection transforms memory information into the shared template space (Fig. 4g, right) and the visual search circuit can be engaged. In this way, dynamically transforming representations may allow the brain to control what and when cognitive computations are engaged.

References for Supplementary Information

1. Miller, E. K., Nieder, A., Freedman, D. J. & Wallis, J. D. Neural correlates of categories and concepts. *Current Opinion in Neurobiology* **13**, 198–203 (2003).
2. Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).
3. Bernardi, S. *et al.* The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell* **183**, 954–967.e21 (2020).
4. Lara, A. H. & Wallis, J. D. Executive control processes underlying multi-item working memory. *Nature Neuroscience* **17**, 876–883 (2014).
5. Lee, H., Simpson, G. V., Logothetis, N. K. & Rainer, G. Phase Locking of Single Neuron Activity to Theta Oscillations during Working Memory in Monkey Extrastriate Visual Cortex. *Neuron* **45**, 147–156 (2005).
6. Buschman, T. J., Siegel, M., Roy, J. E. & Miller, E. K. Neural substrates of cognitive capacity limitations. *PNAS* **108**, 11252–11255 (2011).
7. Masse, N. Y., Yang, G. R., Song, H. F., Wang, X.-J. & Freedman, D. J. Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nat. Neurosci.* **22**, 1159–1167 (2019).
8. Myers, N. E., Stokes, M. G. & Nobre, A. C. Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends Cogn. Sci. (Regul. Ed.)* **21**, 449–461 (2017).
9. Murray, J. D. *et al.* Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *PNAS* **114**, 394–399 (2017).
10. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat Neurosci* **17**, 440–448 (2014).
11. Ruff, D. A. & Cohen, M. R. Simultaneous multi-area recordings suggest that attention improves performance by reshaping stimulus representations. *Nat Neurosci* **22**, 1669–1676 (2019).
12. Miller, E. K. & Cohen, J. D. An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience* **24**, 167–202 (2001).
13. Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D. & Miller, E. K. Synchronous Oscillatory Neural Ensembles for Rules in the Prefrontal Cortex. *Neuron* **76**, 838–846 (2012).
14. Fries, P. Rhythms for Cognition: Communication through Coherence. *Neuron* **88**, 220–235 (2015).
15. Stokes, M. G. *et al.* Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron* **78**, 364–375 (2013).