PNAS RESEARCH ARTICLE | NEUROSCIENCE
PSYCHOLOGICAL AND COGNITIVE SCIENCES

Multiplexing of cognitive encoding by oculomotor networks leads to incidental gaze shifts

Matthew C. Rosen[a,1] and David J. Freedman[a,b,1]

Affiliations are included on p. 10.

Edited by Robert Desimone, Massachusetts Institute of Technology, Cambridge, MA; received October 28, 2024; accepted February 27, 2025

Humans and other animals are adept at learning to perform cognitively demanding behavioral tasks. Neurophysiological recordings in nonhuman primates during such tasks find that the requisite cognitive variables are encoded strongly in core oculomotor brain regions. Here, we assembled a large dataset—11 monkeys performing an abstract visual categorization task, surveyed across more than 1,000 neural recording sessions—to reveal that this produces a robust but uninstructed behavioral "tell," observed in all subjects and experiments: small, cognitively modulated eye movements. We find that these eye movements are causally linked to activity in SC but not LIP, and that they occur following transient alignment of cognitive and saccadic population coding subspaces in SC. This behavioral signature of oculomotor engagement is absent during a similar task that does not require rule-based categorization, suggesting that abstract task behaviors recruit primate oculomotor networks more strongly than previously understood.

working memory | eye movements | parietal cortex | cognition | neurophysiology

Significance

When humans perform working memory tasks, their eye movements disclose a wealth of information about what they are holding in mind. Here, we report a similar behavior in monkeys—small, uninstructed eye movements that reflect task variables in working memory, observed across many subjects and experiments. We find that these eye movements reflect representational cross-contamination within the oculomotor system: neuronal activity in areas that control gaze also strongly encode abstract cognitive variables, even during behavioral tasks which do not require shifts of gaze or attention. This provides insight into how core spatial and oculomotor brain networks are recruited to mediate abstract cognition, and the neural coding principles that underlie multiplexing of multiple task variables by single neural populations.

Humans and other intelligent animals are remarkably adept at learning and performing arbitrary cognitively demanding tasks. Many of our everyday behaviors depend on this capacity for flexibility, both in interpreting incoming sensory stimuli and in mapping decisions to task-appropriate actions. For example, when we shop for groceries, we engage in a general cognitive task of searching for items according to a list, improving in efficiency through experience as we learn the store layout. In doing so, we also solve a series of subtasks: we interpret the sensory features of each item we encounter in light of prior knowledge and present context—e.g., what meals are planned that week—to form a binary judgment ("buy it" or "leave it").

Recent findings show that abstract decision-making tasks recruit brain networks critical for controlling gaze to a surprising degree. Multiple studies have reported stronger encoding of cognitive variables in regions associated with oculomotor function than in the high-order visual cortex (MT, MST) (1, 2) or prefrontal cortex (PFC) (3). Reversible inactivation studies have further validated the causal importance of this encoding for behavior (4, 5). These observations are especially noteworthy because they were made during tasks specifically designed to minimize eye movements, shifts of spatial attention, and differential action planning for different categories, any of which might otherwise explain why canonically oculomotor areas appear to causally support cognitive behavior. Although a substantial behavioral literature has established that (small) eye movements reflect a broad spectrum of cognitive and behavioral states [e.g., spatial attention (6–8), mental fatigue (9–11)], these studies have left open why and how they do so, and what this says about the role that oculomotor areas play in cognition. Given that brain networks for spatial orienting and motor planning are well positioned to influence animals' motor responses, their recruitment during abstract cognitive tasks—particularly when they are extensively trained and highly familiar—could support more efficient or rapid decisions and lessen cognitive load.

The encoding of cognitive variables in neural populations which drive motor actions exemplifies a widespread neural coding challenge with potentially serious behavioral consequences: representational interference. Unless the activity subspaces that encode cognitive vs. oculomotor variables are completely orthogonal, then executing either function could influence encoding of the other, leading to cross-contamination (like a poker player's "tell"). When oculomotor areas encode nonspatial cognitive task variables, for example, such variables might affect eye movements, as has recently been reported in human subjects (12–16). Similarly, ongoing eye movements during task behavior, instructed or otherwise, could interact or interfere with these areas' encoding of cognitive variables. This potential for cross-contamination is reminiscent of that between sensory and mnemonic representations, which recent reports suggest is mitigated by rotation into orthogonal activity subspaces (17).

Author contributions: M.C.R. and D.J.F. designed research; M.C.R. performed research; M.C.R. analyzed data; and M.C.R. and D.J.F. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]To whom correspondence may be addressed. Email: matthew.casciola.rosen@gmail.com or dfreedman@uchicago.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2422331122/-/DCSupplemental.

Published April 8, 2025.

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

Determining how the brain resolves this issue with respect to cognitive encoding in motor areas will be critical for understanding how neural systems reliably control both cognition and action.

To study interference between cognitive and oculomotor representations in the primate brain, we evaluated neural activity and eye movements in many macaque monkeys performing an abstract visual categorization task. High-resolution gaze tracking revealed a clear behavioral index of representational interference in the oculomotor system, present in all subjects and experiments: small, uninstructed eye movements reflecting the category guiding the subject's task performance, despite no instruction or training to do so. This effect diminished significantly during reversible inactivation of SC but not LIP, implicating the SC causally in the emergence of cognitively modulated eye movements. Subsequent investigation of SC activity identified a neural population signature of this "leak": increases in alignment between components of activity encoding categories and upcoming gaze shifts. Behavioral and neural signatures of this leak disappear during a similar task variant that does not require categorization (but which involves similar visual stimuli, task timings, and motor responses), suggesting that recruitment of primate oculomotor circuits for cognition may specifically reflect the information-processing demands of abstraction.

Results

Our work has studied neural representations supporting abstract decisions using a delayed match-to-category (DMC) task paradigm (Fig. 1*A*). During the DMC task, animals group dot-motion stimuli into categories according to a learned category boundary based on motion-direction, and indicate via manual report whether sequences of these stimuli belong to the same category or not. In the delay period (≥1 s) between sample and test stimuli, animals tend to maintain information about the sample stimulus in an abstract format that generalizes across motion directions within a category.

Abstract Categories Are Encoded Especially Strongly in Oculomotor Areas. A meta-analysis of neural recordings conducted from a range of areas during DMC tasks suggests a general trend toward recruitment of brain regions that control gaze and covert attention (Fig. 2*A*). PFC neurons show tuning to learned motion categories (3), as they do to abstract categories informed by other visual features (18) (Fig. 2*G*). However, category tuning in PFC is weaker, emerges more slowly, and is less closely tied to animals' trial-by-trial behavior than in the lateral intraparietal area (3) (LIP), a visual/oculomotor association structure in the dorsal visual pathway (19–21) (Fig. 2*F*). LIP neurons, typified by spatially selective visual and saccadic responses, also show strong, short-latency category selectivity during DMC (1–3, 5, 22–25).

Further work has established that abstract categories are also strongly encoded in other cortical and subcortical areas mediating gaze control and spatial orienting. In areas where these ties are weaker, the corresponding representation of motion categories is weaker: MST, a region just upstream of LIP in the dorsal pathway that also encodes motion directions, shows slightly but significantly weaker motion category tuning (2) (Fig. 2*E*). In MT, a visual area with strong direction-selective responses but only marginal modulation around eye movements beyond that attributable to motion of the retinal image (26), category tuning is absent altogether (1) (Fig. 2*D*). This tendency against tuning to abstract categories in regions primarily engaged in visual feature processing aligns with the relatively weak encoding of abstract categories based on visual objects observed in the IT cortex (27). In contrast, categories are strongly encoded in the superior colliculus (SC), a subcortical structure critical for spatial orienting behaviors, especially saccades and spatial attention (5) (Fig. 2*H*). Additional recordings in the medial intraparietal area (MIP), which is associated with hand movements, indicate that this tendency may be concentrated on regions involved in oculomotor control, as category encoding was significantly stronger and emerged with a significantly shorter latency in LIP than in MIP (28).

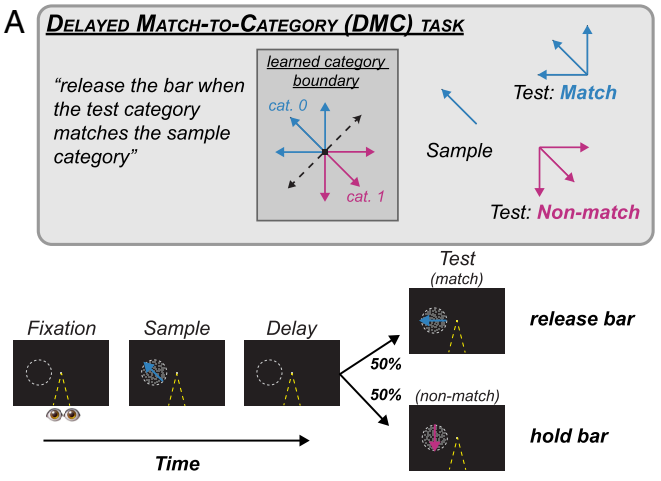

Fig. 1. Delayed match-to-category task. (*A*) *Top*: Description of DMC task logic and stimulus-category assignment. *Bottom*: Schematic depicting timing of task events during DMC. Note that on non-match trials, the non-matching test stimulus was followed by a brief delay and a second test-stimulus which always matched the category of the sample (and required a bar release). Specific task epoch times differed by experiment (*SI Appendix*, Table S1).

Small Eye Movements Reflect Abstract Categories. We next sought to evaluate the degree of representational cross-contamination between abstract categories and eye movements. If the encoding of categories observed in oculomotor areas interferes with those areas' encoding of eye movements, then eye positions should differ systematically by category. The only explicit constraint on these differences, if they exist, is that they must be small, as trials terminate when subjects' gaze moves outside a small central fixation window (~2 degrees radius visual angle). To determine whether such an uninstructed relationship between categories and eye movements is a general feature of behavior during the DMC task, we assembled a large set of behavioral and physiological recordings—18 datasets collected from 11 unique monkeys, including 298,492 correctly completed trials across 1,063 recording sessions (*SI Appendix*, Table S2). This afforded us a unique opportunity to evaluate the prevalence of category-modulated eye movements in a setting with sufficient power to detect a potentially small effect, and across variability in individual animals and studies.

Visualization of eye movements during the memory delay on single trials (example session, late delay, Fig. 3*C*) revealed separation of gaze position according to the remembered sample stimulus category. These patterns did not exhibit a stereotyped relationship with motion direction across subjects or experiments, and were generally idiosyncratic, differing in onset time, magnitude, and direction (one example session from each of four different subjects, Fig. 3 *A* and *B*). To capture the degree of separation in eye positions by category quantitatively, we fit linear SVM classifiers to predict category identity from gaze positions in each behavioral session (Fig. 3 *D* and *E*; *Materials and Methods*). In all datasets, category decoding accuracy from gaze position was significantly greater than chance (for each dataset, Wilcoxon signed-rank test, data vs. label-shuffle control, one-sided, with N = # sessions; FDR controlled via Benjamini–Hochberg procedure, all $P < 0.001$).

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

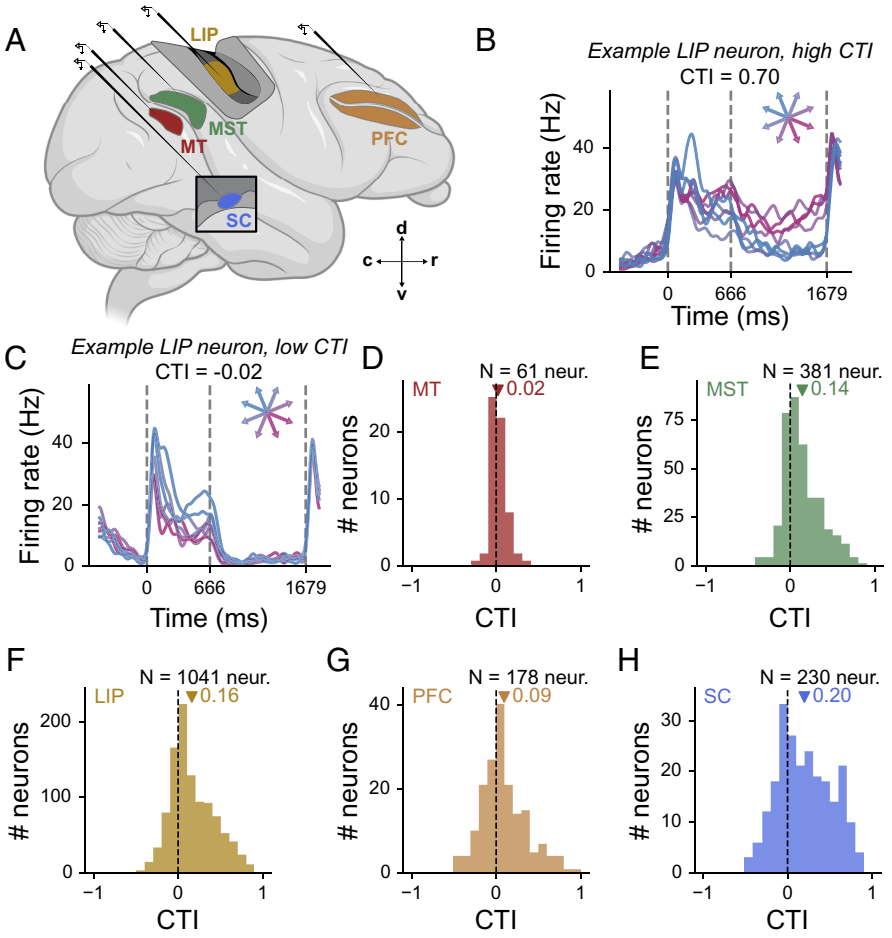

Fig. 2. Abstract categories are strongly encoded in areas associated with gaze control. (*A*) Schematic of the brain with recorded areas highlighted. (*B*) Firing rate of one example LIP unit with high category tuning index (CTI) at the end of the delay, averaged by sample direction and colored by category. (*C*) Same as *B*, but for an example unit with low CTI. (*D–H*) Histogram of CTI of recorded neurons from MT, MST, LIP, PFC, and SC. Labels indicate mean CTI for all units recorded in each respective area.

To assess whether these gaze differences reflected the category of the sample stimulus per se, rather than the elementary visual feature informing category (motion direction), we also fit classifiers to decode counterfactual category labels from eye positions—labels assigned according to a boundary orthogonal to the true instructed boundary (Fig. 3 *D* and *E*; *Materials and Methods*). If gaze positions reflect the sample motion direction, then counterfactual and true category labels should be equivalently decodable. If gaze specifically reflects the learned categories, however, then labels assigned according to an orthogonal category boundary should be decoded less accurately than the true labels. Consistent with visualization of gaze traces by direction (Fig. 3 *A* and *B* and *SI Appendix*, Fig. S1), we found evidence consistent with the latter possibility: in all datasets, counterfactual labels were decoded significantly less accurately than true labels (for each dataset, Wilcoxon signed-rank test, data vs. counterfactual label control, one-sided, N = # sessions in dataset; all $P < 0.001$), suggesting that separation in eye movements during DMC specifically reflects the category identity of the sample motion stimulus.

We next sought to determine how closely these differences in gaze reflect the decision variable guiding behavior. This value is correlated with but not identical to the presented category since the animals do make errors, particularly on difficult trials with motion stimuli close to the category boundary. If eye movements track the decision variable, then they should reflect the presented category less faithfully on error trials than on correct trials. Eye movements on error trials were generally dissimilar to those on correct trials of the same presented category (Fig. 3*F*). We developed a congruence index (CI), which measures how closely eye traces match those on correct trials of the presented vs. the opposite category (*Materials and Methods*). Across all datasets, CI at the end of the delay was significantly lower on error trials than on held-out correct trials (Fig. 3 *G* and *H*; one-sided Wilcoxon signed-rank test, CI on error vs. correct trials, N = 18 datasets, $P < 0.001$), consistent with the interpretation that differences in eye movements by category reflect the abstract decision variable guiding the animals' match/nonmatch judgments. Together, these results suggest that uninstructed eye movements reflect the abstract cognitive variable guiding behavioral performance, and constitute a general feature of animals' behavior during the DMC task.

Category-Aligned Gaze Reflects Phasic Modulations of Microsaccade Rate and Amplitude. Category-specific gaze patterns could arise through any of several qualitatively different ways—for example, saccades executed in systematically different directions on trials with different categories, or at systematically different times, or of systematically different amplitudes. These differences in gaze position could also reflect modulation of the speed or direction of fixational drift.

To resolve more specifically which aspects of eye movements contribute to the category-specific gaze patterns we report, we evaluated microsaccades and fixational drift on single trials within each behavioral session (*SI Appendix*, Fig. S2*A*). We focused this analysis on a 1,500 ms period aligned to the beginning of the delay

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

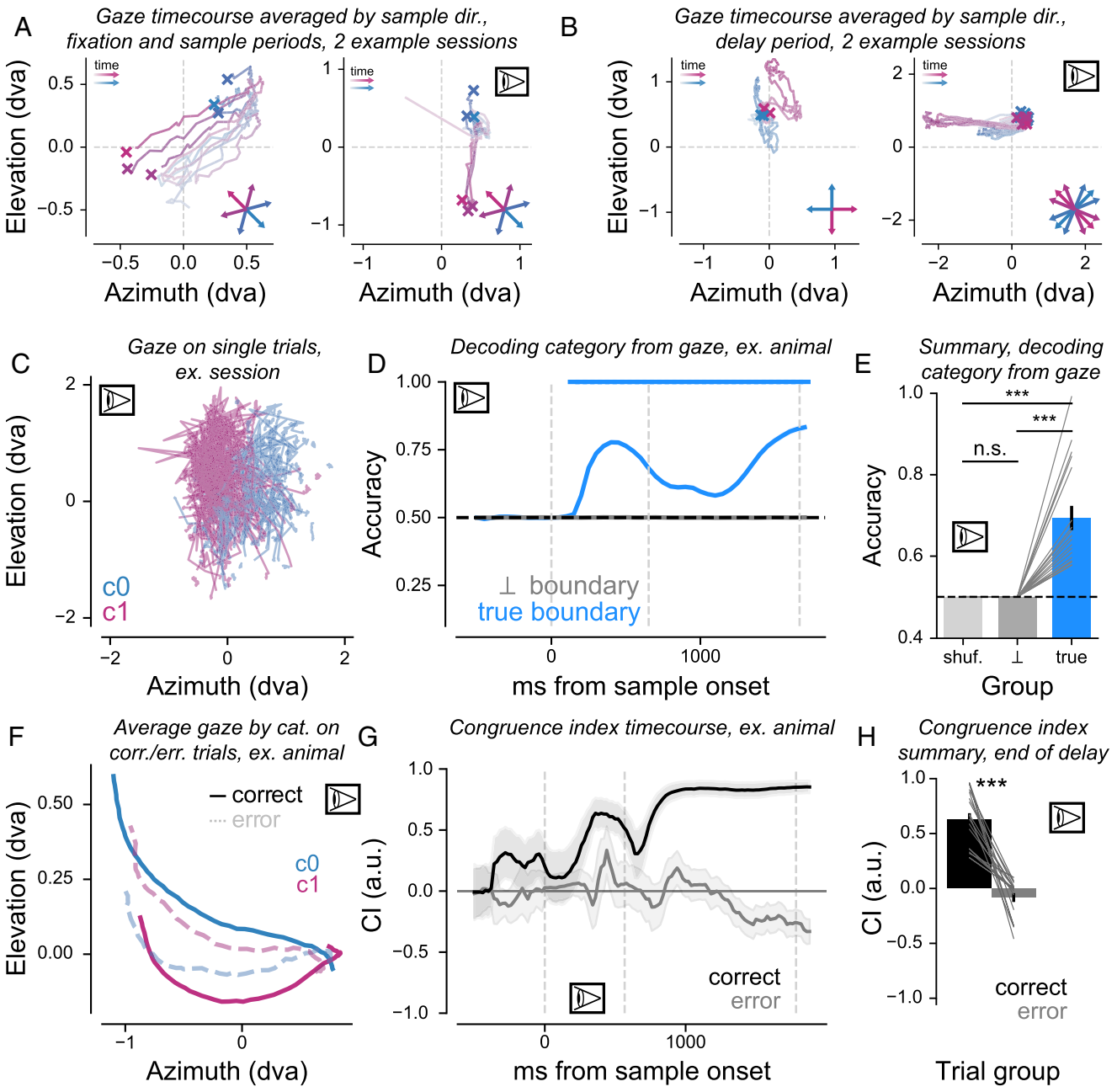

Fig. 3. Small eye movements reflect abstract categories. (*A* and *B*) Eye position time-courses, averaged by sample direction, for four different animals. Colors corresponding with different sample directions are shown in the schematic at the *Bottom Right* of each panel, and "x" markers denote the time of sample stimulus offset. Increases in saturation reflect the temporal order of the timecourse (early = less saturated, late = more saturated). (*A*) Eye position time-courses for two example animals, from fixation acquisition to sample offset. (*B*) Eye position time-courses for two different example animals, from sample offset to the end of the delay. (*C*) Eye movements on all single trials from one example session, Monkey J, final 250 ms of the delay period. (*D*) Category decoding from eye positions through time, Monkey J (mean ± SEM across sessions, N = 109; true labels: blue, counterfactual labels: gray). Blue lines at top denote moments where classifier performance significantly exceeded the chance level of 50% (paired *t* test at each timepoint, N = 109 sessions, $P < 0.05$). (*E*) Summary of classifier performance across datasets (mean ± SEM, N = 18) relative to shuffle control (light gray), counterfactual label control (dark gray). (*F*) Monkey S, average delay-period eye position traces colored by sample category, correct trials (solid), and error trials (dashed). (*G*) Congruence index time-course on error trials and held-out correct trials, Monkey S. Error bars denote SEM estimated via bootstrap. (*H*) Summary of congruence index at the end of the delay period across datasets, mean ± SEM (N = 18).

period (e.g., encompassing 600 ms of sample presentation and 900 ms of the delay), chosen because it broadly encompassed the emergence of category-specific gaze patterns across all datasets. In this period, we analyzed the distribution of five eye movement features for each category (*SI Appendix*, Fig. S2*A*): the direction and speed of fixational drift (*SI Appendix*, Fig. S2 *B* and *C*), and the direction, amplitude, and frequency of microsaccades (*SI Appendix*, Fig. S2 *D–F*). Because each of these parameters can vary through time, and when modulated (even transiently) can lead to persistent changes in gaze position, we evaluated the difference between categories for each eye movement feature in successive 150 ms bins. This approach allowed us to capture, for each behavioral session, the moments at which these eye movement parameters differed significantly between the two categories (*Materials and Methods*).

We detected significant differences between categories in these various eye movement parameters, at least transiently, in most individual sessions (714/1,063 sessions with significant difference in at least one feature for at least one 150 ms bin; *Materials and Methods*). These differences were generally phasic, rarely persisting

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

for more than 300 ms at a time, and idiosyncratic (different animals showed differences for different features at different times). The most frequent effect we observed was a significant difference in the rate of microsaccades between categories (650/1,063 sessions, or 61.1%; *SI Appendix*, Fig. S2*D* and *Materials and Methods*). In just under half of sessions, the speed of fixational drift and the amplitude of microsaccades differed significantly between categories (*SI Appendix*, Fig. S2 *C* and *F*; drift speed: 465/1,063 sessions, or 43.7%; microsaccade amplitude: 485/1,063 sessions, or 45.6%). Significant differences between categories in the direction of microsaccades or fixational drift were observed, but generally less frequently (*SI Appendix*, Fig. S2 *B* and *E*; microsaccade direction: 390/1,063 sessions, or 36.7%; drift direction: 256/1,063 sessions, or 24.1%). These results suggest that category-specific gaze patterns can arise through transient differences between categories in a variety of eye movement features, but that they most frequently reflect a tendency to make more microsaccades on trials of one category than trials of the other.

Category-Correlated Eye Movements Have Little Effect on Category Encoding in MST, LIP, and SC. Given this robust reflection of categories in gaze patterns, we set out to study the relationship between eye movements and the neural encoding of categories in oculomotor areas. We first evaluated how, if at all, the neural encoding of remembered categories covaries with the execution of category-correlated eye movements. Even a small gaze shift can drive response modulations across the visual system, whether through changes to the retinal image (29), proprioceptive inputs, efference copies, or other means (*SI Appendix*, Fig. S3). Such shift-driven responses could in principle contribute to or interfere with the neural population encoding of categories that we have previously reported.

We investigated these possibilities in MST, LIP, and SC by training category decoders on pseudopopulation activity assembled from the delay period of trials where category-correlated eye movements were the most vs. least prominent (*SI Appendix*, Fig. S4*A*; *Materials and Methods*). If category-correlated gaze shifts drive responses that interfere with category encoding, then decoders trained on trials with more separation of eye movements by category ("Hi") should be less effective than those trained on trials with less separation of eye movements by category ("Lo") (*SI Appendix*, Fig. S4 *B*, *Left*). Conversely, if these responses instead contribute to the neural encoding of categories, then decoders trained on "Hi" should be more effective than those trained on "Lo" (*SI Appendix*, Fig. S4 *B*, *Right*).

However, we generally found no significant difference between category decoding from "Lo" vs. from "Hi" (*SI Appendix*, Fig. S4*C*; MST: significant difference in 0 of 2 datasets; LIP: significant difference in 2 of 18 datasets, SC: 0 of 2 datasets; MST: 0 of 2 datasets; for each dataset, bootstrap test, $P < 0.05$). This null effect was not due to a failure of our procedure to divide trials by the extent to which they contained category-specific gaze patterns, as eye positions on trials of different categories in "Hi" were significantly farther apart on average than in "Lo" for all datasets in both LIP and SC (for each dataset in each area, significance of difference in category-correlated eye separation assessed via Wilcoxon signed-rank test, one-sided, all $P < 0.01$, N = # sessions per test; FDR controlled via Benjamini–Hochberg procedure). In the two datasets for which we detected a statistically significant difference in decoding, the difference was small, around 10%, but positively skewed (higher when eye movements reflected categories more strongly). Because all decoding comparisons were performed within individual datasets, the difference in decoding observed in these two datasets could not be attributed to different task variants, animals, neurons, or trial counts, demonstrating that small category-correlated eye movements can in principle contribute to measurable increases in category encoding. The failure to detect this effect across most datasets under similar conditions, however, suggests that category encoding in these areas does not just reflect category-correlated eye movements.

Reversible Inactivation of SC But Not LIP Diminishes Category-Correlated Eye Movements. We next sought to determine how category encoding across our measurements of the oculomotor network in LIP and SC may relate to category-correlated eye movements. To do so, we evaluated eye position records during inactivation experiments using the GABA-A agonist muscimol (4, 5), which significantly impaired DMC choice behavior. Such impairments, consistent with these areas playing a causal role in the monkeys' categorical decisions, were especially strong during SC inactivation (average change in DMC task accuracy during SC inactivation relative to control sessions: Monkey N, −28% accuracy; Monkey S, −24%; average change in DMC task accuracy during LIP inactivation relative to control behavior: Monkey M: −5% accuracy; Monkey Q, −7%; see *Materials and Methods*).

In addition to these substantial deficits in animals' match/nonmatch decisions based on the learned motion categories, small eye movements were also impacted substantially during SC inactivation (see *Materials and Methods* and *SI Appendix*, Fig. S5 *A* and *B*; Monkey Sta: significant difference from control in 7/7 sessions, 2-sample Epps-Singleton test, $P < 0.01$; Monkey N: 6/6 sessions), and reflected the presented category significantly less faithfully than during control behavior (Fig. 4 *A* and *C*; Wilcoxon rank-sum test, Monkey Sta: $P < 0.001$, N = 7 sessions; Monkey N: $P = 0.0024$, N = 6 sessions). In contrast, small eye movements were minimally impacted during LIP inactivation (*SI Appendix*, Fig. S5 *C* and *D*), and did not reflect the presented category significantly differently during the delay period relative to control behavior (Fig. 4 *B* and *C*; Wilcoxon rank-sum test, animal Ma: $P = 0.22$; animal Q: $P = 0.74$). These differences were reflected in a large reduction in category classification from eye movements during inactivation of SC relative to control (Fig. 4*D*, blue points) and a comparatively smaller effect during inactivation of LIP (Fig. 4*D*, yellow points). Note that the time-courses during control and inactivated behavior shown in Fig. 4 *A* and *B* reflect data from different animals during inactivation of SC (Fig. 4*A*) and LIP (Fig. 4*B*), for whom the time-course and amplitude of category-correlated eye movements were generally idiosyncratic.

SC's Population Encoding of Categories and Gaze Shifts Become Aligned Before Category-Correlated Eye Movements. Motivated by the collapse of category-correlated eye movements during inactivation of SC, we turned to SC activity to resolve how such eye movements may arise. Consistent with prior work (30), we found that single SC neurons were significantly modulated in the leadup to small gaze shifts, and that SC population activity supports decoding of these shifts' directions (*SI Appendix*, Fig. S6). Together with our previous finding of strong category-selective activity in the SC, this pointed to a candidate explanation for how category-modulated eye movements may emerge: through leak of category encoding into the components of SC activity that contribute to gaze shifts. If so, then the neural population encoding of categories and gaze shifts in SC should display moments of nontrivial alignment with one another. Further, to be capable of driving category-correlated eye movements, this alignment must emerge before category-correlated eye movements do.

To test these predictions, we analyzed the relationship between the components of SC activity that encode categories vs. gaze shifts

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

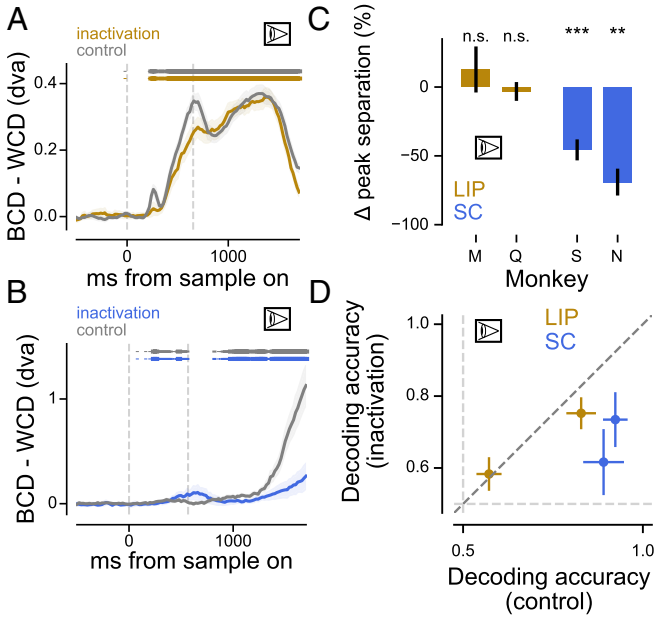

Fig. 4. Reversible inactivation of SC but not LIP diminishes category-correlated eye movements. (*A*) Time-course of separation in eye movements according to category during LIP inactivations, Monkey Q; inactivation behavior in yellow and control behavior in gray. Error bars denote SEM across sessions. (*B*) Same as *A*, but for SC inactivations in Monkey N, with inactivation behavior in blue. (*C*) Percent change in peak eye-movement separation by category during inactivation sessions relative to control sessions, LIP and SC. (*D*) Decoding of category from eye movements during inactivation vs. control sessions; each point represents data from one individual animal, colored by area inactivated (yellow: LIP; blue: SC). Error bars represent SD across sessions.

(Fig. 5 *A* and *B*). Using activity during a separate memory-guided saccade task (MGS), we identified a *shift-coding direction* **sd**—the direction in N-dimensional population activity space containing the most variation in the leadup to eye movements in different directions (*Materials and Methods*). We used the same neurons' activity during DMC to identify a *category-coding direction* **cd**—the direction in population activity space that best discriminates between categories (Fig. 5*A*). The extent to which the neural population encoding of gaze shifts is aligned or orthogonal to its encoding of categories is captured by θ, the acute angle between **sd** and **cd** (Fig. 5*B*). We evaluated whether **cd** is significantly more aligned to **sd** than expected by chance by comparing θ to θ_{shuf}, the distribution of angles between **sd** and **cd** computed from randomized category labels (*Materials and Methods*).

This analytical approach revealed significant alignment between **sd** and **cd** immediately before the emergence of category-correlated eye movements (Fig. 5*C*). For each subject with SC recordings, we identified the latency to category-correlated eye movements (Monkey Sta: 330 ms after sample onset; Monkey N: 970 ms after sample onset; *Materials and Methods*), and used activity in the preceding 200 ms to compute **cd**, θ, and θ_{shuf}. For both animals, θ was significantly lower than θ_{shuf}, indicating alignment of **cd** and **sd** before gaze reflected categories (Fig. 5*C*, blue points; bootstrap test, one-sided, both $P < 0.05$). This suggests that category encoding in SC, particularly in activity components predictive of small upcoming gaze shifts, may drive category-correlated eye movements.

Oculomotor Areas Are Specifically Recruited by Tasks Requiring Abstract Cognition, Not Just Matching or Short-Term Memory. Our findings so far demonstrate that oculomotor areas encode abstract categories during DMC–which we propose leads to small, category-correlated eye movements. However, we seek to understand which of DMC's numerous task demands (direction discrimination, categorization, working memory, match/nonmatch decisions) best explains the incidence of these gaze patterns. To address this question, we examined data collected during a similar delayed match-to-sample-direction (DMS) task (Fig. 6*A*), which uses identical stimuli, timings, and manual motor responses to DMC but without the demand to categorize the motion stimuli. If oculomotor regions are generally recruited for tasks requiring animals to discriminate motion, match stimuli, or hold information in their short-term memory, demands which DMC shares with DMS, then the behavioral and neural signatures of oculomotor involvement in DMC should also appear during DMS. If oculomotor regions' engagement reflects computational demands for abstraction, however, then these signatures should not appear during DMS.

To determine whether oculomotor networks encode task variables during DMS, we analyzed datasets from four animals performing the DMS task, including eye movement measurements in all four animals and neuronal recordings from LIP in 2 (Monkeys D and He) (24, 31). In these two animals, LIP recordings were obtained at three sequential training phases: 1) after learning to perform the DMS task; 2) at an intermediate stage of learning to perform the DMC task ("DMCearly"); and 3) after extensive additional training to perform the DMC task. This allowed us to directly compare neural activity and behavior in the same animals across tasks.

Unlike during DMC, we found no evidence, neural or behavioral, that oculomotor networks are engaged in encoding task variables during DMS. In all four datasets with measurements of eye movements during DMS task behavior, eye movements did not differentiate significantly between different sample stimulus directions, especially during the memory delay (Fig. 6*B*). Neural activity

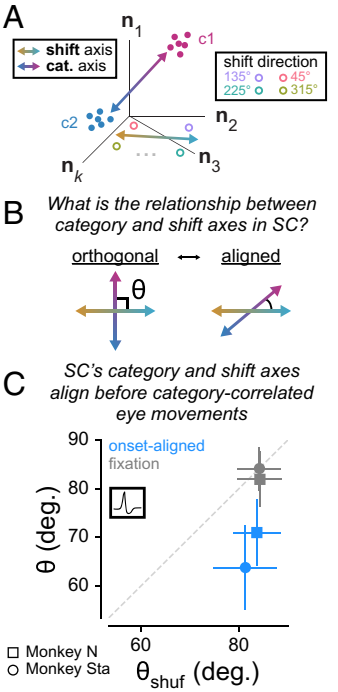

Fig. 5. SC's population encoding of categories and gaze shifts become aligned before category-correlated eye movements. (*A*) Analysis schematic: identifying shift- and category-coding directions in population activity space. (*B*) Graphical depiction of the key quantification: what is the angle θ between the shift-coding and category-coding axes in SC? (*C*) Angle between SC's shift-coding axis and its category-coding axis, computed using true (ordinate) vs. shuffled (abscissa) category labels. Blue: category-coding axis computed in the 200 ms window leading up to the emergence of category-correlated eye movements ("onset-aligned"); gray: category-coding axis computed using activity during the fixation period before sample stimulus onset. Square markers: Monkey N; circular markers: Monkey Sta. Error bars denote SEM (estimated via bootstrap). The gray dashed line denotes unity.

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

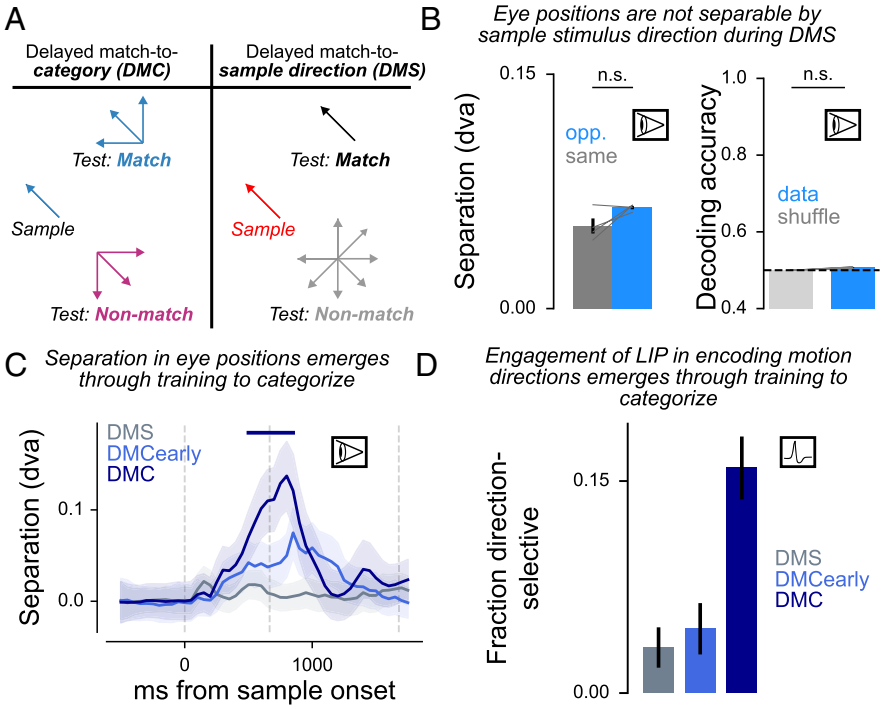

Fig. 6. Oculomotor areas are specifically recruited by tasks requiring abstract cognition, not just matching or short-term memory. (*A*) Schematic highlighting the key distinction between DMC and the delayed match-to-sample task: subjects need to indicate whether sample and test directions match identically, instead of whether they belong to the same learned category. (*B*) *Left*: Peak delay-period separation in eye movements for DMS trials with the opposite (light blue) or same (gray) sample direction, d.v.a.; bars show mean ± SEM across animals performing DMS (N = 4), while gray lines show data for individual animals. *Right*: Average decoding of sample direction dichotomies during DMS (*Materials and Methods*) from true and shuffled motion direction labels. As before, gray lines show data for individual animals (all near 50%), and error bars (extremely small) indicate ± SEM. (*C*) Monkey D, time-course of difference in eye movements for trials with the opposite vs. same sample direction, DMS (gray), early in DMC training (light blue), after learning DMC (dark blue). Timepoints of significant difference from 0 (bootstrap, $P < 0.05$) shown for each task in the correspondingly colored line at the top of the figure. (*D*) Fraction of units discriminating between different sample directions during the late delay period during each task. Significance of selectivity neurons determined by the Kruskal–Wallis H test, $P < 0.001$ (FDR controlled using the Benjamini–Hochberg procedure).

in LIP during the DMS delay also encoded the sample direction only weakly, even though the animals' behavior indicated that they successfully remembered the direction of the sample stimuli (Fig. 6*D*). In the two animals with LIP recordings during both DMC and DMS (Monkeys D and He), the fraction of neurons with responses differentiating between different sample directions through the delay was significantly higher during DMC than DMS (Fig. 6*D*; 2-proportion z-test, one-sided, $P < 0.001$), and eye movements during the late sample/early delay period reflected remembered stimuli significantly more strongly (Fig. 6*C*; bootstrap, $P < 0.05$).

Measurements of activity and behavior during the intermediate DMC training stage (DMCearly) indicate further that the extent to which oculomotor areas are engaged during DMC reflects the extent to which the animal employs an abstract strategy to categorize the stimuli. At this intermediate stage, when the animals' performance was above chance but suggested a similarity-based strategy (i.e., comparing the angular difference between sample and test stimuli) rather than an abstract categorical strategy (24, 32), neural activity in LIP and eye movements both reflected the assigned category of the sample stimulus only weakly (Fig. 6 *C* and *D*). These observations together suggest that the oculomotor system may be selectively engaged by DMC's demand to group stimuli into abstract categories.

Discussion

The ability to learn and perform arbitrary cognitively demanding tasks is fundamental for intelligent behavior. Increasing evidence points to a surprisingly central role for oculomotor areas in encoding the cognitive variables supporting performance of these tasks. Here, we studied how the nervous system multiplexes these distinct functions in the same circuits through a behavioral signature of their interaction—gaze patterns that reflect task variables. We validated the robustness of this uninstructed feature of behavior by assembling a large corpus of single-unit electrophysiological and behavioral data collected from many monkeys performing the same abstract categorization task. In this dataset, we found that neural representations of cognitive task variables are concentrated in brain regions involved in oculomotor control, such as LIP and SC, even though the animals in our studies were required to maintain central fixation throughout each trial and reported their decisions via a movement of their hand and not their eyes.

A previous study from our group found that neuronal population encoding of categories and eye movements in SC inhabit near-orthogonal representational subspaces (5). We speculated that the category-correlated gaze shifts observed here might arise from incidental leak from the category subspace into the gaze-shift subspace. Indeed, we found a particularly strong link between SC activity and the emergence of category-modulated eye movements: when SC was temporarily inactivated, these eye movements diminished, and before category-correlated eye movements emerge, the principal category-coding direction in SC population activity is aligned with the principal gaze-shift-coding direction.

Our behavioral findings resemble a similar phenomenon reported in humans performing cognitive tasks that demand working memory (12–16). Despite explicit verbal instructions not to deviate from central fixation, human subjects' small fixational eye movements disclose remembered information about task variables, including features that are both relevant (16) and irrelevant (14, 15) for the instructed behavioral report. These eye movements have been

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

proposed to reflect engagement of the brain's gaze control system in selecting information from short-term memory as subjects prepare their behavioral response. Here, we identify the physiological underpinnings of a similar uninstructed feature of behavior in the macaque: neural representations of abstract cognitive task variables in core oculomotor regions that occupy population activity modes which are nontrivially aligned to those driving eye movements.

Our physiological findings also echo reports of widespread modulations of neural activity by small uninstructed motor actions (33–35). Consistent with prior reports (26, 36–40), many units across the regions from which we recorded showed significant activity modulations after these movements' onset (*SI Appendix*, Fig. S3). Uncovering the source of these modulations–whether they reflect a visual response, proprioceptive input, a result of efference copies signaling the movement command broadly across the cortex, or another influence altogether—remains an important topic for future work.

The potential of such fixational eye movements to result in neural activity modulations also raises an important question: what, if any, is the functional importance of small, category-correlated eye movements? At one extreme, such eye movements might reflect a strategy that animals adopt to help meet the mnemonic demands of the DMC task. According to this class of possibility, category-correlated eye movements would effectively act to externalize the value of the "internal" cognitive variable that the animal needs to remember. This process could facilitate the maintenance of information about categories in working memory, either directly [category-correlated movements driving category-correlated neural responses, as discussed above, or helping to refresh them (41)] or indirectly (by positioning gaze differently with respect to the test stimulus, depending on the category of the sample stimulus). At the other extreme, these gaze patterns may just reflect engagement of circuits mediating oculomotor control in computing or representing learned categories, without the eye movements themselves serving any functional role. Even if category-correlated eye movements drive category-correlated neural responses, such responses do not necessarily furnish a signal that is read out to support task behavior.

What determines the extent to which gaze control networks represent cognitive variables during task performance? During the delayed match-to-sample task, which was identical to the categorization task in nearly all respects but lacked the abstract categorization demands, we did not observe gaze-shifts which reflected the animals' decisions or the contents of their working memory. This suggests that modulation of gaze shifts by task variables is particularly strongly associated with abstract cognitive behavioral demands. However, these demands, which were shared across all DMC task variants, nonetheless evoked effects of varied magnitudes across subjects. We propose that this variation, and engagement of the oculomotor system during a cognitive task more broadly, reflects the extent to which a subject relies on an abstract strategy for solving that task. This view is consistent with our failure to observe neural or behavioral signatures of oculomotor systems' engagement during either the DMS task or during mid-training stages of the DMC task, when animals relied on a similarity-based strategy (Fig. 6). It also aligns with the particularly strong effects we observe in the DMC task variant using two orthogonal boundaries (*SI Appendix*, Fig. S7), which is most likely to be solved using an abstract strategy since it discourages similarity-based, nonabstract strategies. Future work conducting longitudinal recordings across the long-term learning process will be especially interesting for comparing neural encoding as a function of training, expertise, and behavioral strategy.

Why might abstract cognitive tasks like DMC engage brain regions that mediate oculomotor control? These areas comprise an evolutionarily ancient system for orienting toward salient spatial locations in the environment–critical for survival behaviors like acquiring food and avoiding predators. This system is thought to operate by organizing information about the external world according to its location in egocentric coordinates, forming a kind of "location-addressable" repository. Our observation that oculomotor areas mediate abstract cognitive computations may reflect the flexible extension of this repository to organize information that is not necessarily spatial, perhaps via projection into a covert spatial representation. The engagement of this system in cognitive tasks through learning may also reflect the degree of volitional control that subjects are able to exert over neural activity in oculomotor areas. On a whim, primates can rapidly and flexibly shift their gaze or attention—both externally, e.g., within their visual field, and internally, with respect to information they hold in memory. These shifts, which can occur even without an overt movement, signal a generally computationally useful capacity to permit rapid selection and control of information at will. The neuronal networks driving these shifts may thus provide a rich, malleable substrate for the volitional control over internal representations that occurs during abstract cognition.

Materials and Methods

Details of all reanalyzed datasets have been published previously (1–3, 5, 22–25, 28); key differences between them are emphasized below.

Subjects. Behavioral and neural data were collected from 11 adult male rhesus macaques (*Macaca mulatta*). All procedures involved in animal care complied with NIH guidelines and policies, as well as those of Harvard Medical School [datasets from Freedman and Assad (1), "FA06," and Fitzgerald et al. (22), "FFA11"] and the University of Chicago IACUC (all other datasets).

Behavioral Tasks. Animals performed all behavioral tasks while head-restrained and seated inside a custom made primate chair. Eye positions were monitored continuously during task performance either by scleral search coil (datasets FA06 and FFA11; sampling rate = 200 Hz, Riverbend Instruments), or optically via a video-based eye-tracking system (Eyelink 1000; SR Research, Ottawa, Canada; sampling rate = 1,000 Hz) (all other datasets). Unless specified otherwise, all motion stimuli were 100% coherence, high-contrast dot-motion stimuli.

Delayed match-to-category (DMC) task. Briefly, the DMC task required animals to indicate whether sequentially presented visual motion stimuli belonged to the same learned abstract category as an initial sample motion stimulus presented at the beginning of each trial. Trials began when the animal grasped a manual touch-sensitive bar and fixated a central spot (~0.2 dva). Animals needed to maintain fixation and hold the bar for the duration of the trial until a test stimulus belonging to the same category as the sample was presented; once this occurred, they needed to release the bar to receive a fluid reward. Animals received extensive training to perform the DMC task, and their task performance saturated above ~85% correct on completed trials.

Although the sequence of events was identical in all experiments in which animals performed the DMC task, the durations of the task epochs differed slightly between experiments, as did the set of motion stimuli used and the category boundary. A full table of sample motion directions used for each experiment (and their mappings to each category) is shown in *SI Appendix*, Table S1, along with the exact timings used for each epoch and additional details about the dot-motion stimuli.

Memory-guided saccade (MGS) task. In most recording sessions, animals also performed a MGS. After the animals fixated a central spot (~0.2 dva) for 500 ms, a target was briefly flashed (100 to 300 ms) at 1 of 8 isoeccentric locations evenly spaced around the fixation spot (typical eccentricity: 5 to 10 dva). After a brief delay (~1 s), the fixation point was extinguished, and the animals needed to make a saccade to the previously shown target location.

Surgical Procedures and Electrophysiological Recordings. All surgical and recording procedures for each previously published dataset are described in detail in the associated publication. In all cases, recording chamber placement was guided by MRI.

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

Recordings were conducted from single tungsten electrodes (FHC Inc.) or linear electrode arrays (Plexon; interelectrode spacing ≥ 50 μm, 16 to 32 channels), lowered through a grid (Crist Instruments or NAN instruments) in dura-piercing guide tubes via a microdrive system (NAN Instruments, David Kopf Instruments). Recorded signals were spike-sorted offline (Plexon, Kilosort) except for recordings conducted in the Assad lab which were sorted online with a dual window discriminator (Bak Electronics). Further technical details about recordings, including recording locations and chamber placement, can be found in the relevant previous publication for each experiment. Relevant information about dataset shape (number recording sessions per brain area, number units per brain area, mean number of correct trials per session) for all datasets is summarized in *SI Appendix*, Table S2.

SC recordings and laminar position. Although SC recordings were targeted primarily to the superficial and intermediate layers, laminar position was not taken into account for the analyses described in this manuscript.

Stimulus placement. For all datasets where recordings were conducted with single electrodes rather than electrode arrays, stimuli were placed inside LIP neurons' receptive fields (RFs) to drive reliable responses. In the more recent experiments, where recordings were obtained using linear electrode arrays (Plexon V- or S-probes), stimuli were placed within the estimated RF of one of the currently recorded units. Where relevant, visual RFs were identified online using responses during MGS trials performed at the beginning of the session, and in some cases also with a flash-mapping task (23, 42).

Muscimol inactivation. All details for both experiments in which muscimol was used to perturb neural activity during the DMC task are included in the relevant publications (4, 5), including muscimol concentration/volume for each session. Importantly, the larger deficit in DMC choice accuracy during SC vs. LIP inactivation cannot be explained by the concentration or volume of muscimol used alone, as substantially larger volumes and concentrations were used in LIP inactivation experiments.

Data Analysis. All analyses were performed using custom software written in Python v3.8. Unless specified otherwise, only behavior/neural activity from correct, completed trials were used. When assessing significance during comparisons between time series (e.g., decoding accuracy relative to chance, or compared between areas), false discovery rate was controlled using the Benjamini–Hochberg correction.

Category tuning index (CTI). To quantify selectivity for the learned category of a motion stimulus, we used the CTI described in ref. 1. Briefly, the index is given by the ratio of (BCD − WCD)/(BCD + WCD), where BCD is the absolute difference in firing rate for pairs of directions in the different categories and WCD for pairs of directions in the same category. To ensure that BCD and WCD were computed from a balanced set of angular differences, we only used pairs of directions separated by between 0° and 180°. For each unique angular difference in this range, we computed both BCD and WCD, averaging each separately across angular differences to obtain a single BCD and WCD value for each unit. These averaged BCD/WCD values were used to compute the CTI, whose values ranged from −1 to 1. Values near −1 indicate different responses to directions in the same category and similar responses to directions in different categories, while values near +1 indicate similar responses to directions in the same category and different responses to directions in different categories.

Quantifying separation in eye positions by category/sample direction. To quantify the time-course of separation in eye positions by sample stimulus direction (Fig. 6*C*) or category (Fig. 4 *A* and *B*), we 1) averaged eye position traces for groups of trials of each unique sample direction; 2) computed the difference between these averages through time for each pair of sample directions; and 3) separately averaged the resulting time-courses for pairs of directions belonging to the same (different) category to generate a within-category (between-category) time-course of difference in average eye position. Trial groups for each unique sample direction were subsampled from the full set of correct completed trials across all sessions of behavior for each animal. N = 200 trials were subsampled for each motion direction, with qualitatively similar results obtained for other choices of N. To facilitate combining eye movement data across different sessions, we mean-subtracted each session's data before grouping, subsampling, and averaging. We used bootstrapping (repeatedly subsampling new groups of trials with replacement for each motion direction, N = 1,000 times) to estimate the SEM separation in eye positions by category and to assess the significance of the difference between the BCD and WCD.

An analogous procedure was performed to compare separation in eye movements according to sample motion direction across tasks (Fig. 6). As above, we subsampled groups of trials for each unique sample direction. However, we only computed difference time-courses for pairs of directions separated by 180°—pairings of directions which the animals would need to treat as nonmatching, whether they were performing DMS or DMC. Averaging the difference trace across every such pairing of directions yielded an opposite-direction difference time-course, analogous to the between-category difference. To assess the level of difference in eye movements on trials with the same direction, we additionally partitioned trials for each motion direction into two nonoverlapping groups, and subsampled (with replacement) 200 trials from each group for each direction. Same-direction differences were estimated by the difference in average eye position traces within the two separate groups sharing the same sample motion direction.

Classification of category from eye positions. To capture the extent to which eye movements within single sessions/on single trials reflected the presented stimulus category, we used a classification-based approach. For each behavioral session, we fit a linear SVM classifier to predict stimulus category from eye movements. Classifier performance was cross-validated by splitting trials belonging to each category in half for each session, one half for training the classifier and the other for testing. Trial group sizes were equated for each category by subsampling before training/testing. This procedure (randomly split trial groups in half for each category to determine train/test sets) was repeated 1,000 times to obtain a distribution of cross-validated classification scores for each session. We report the mean of this distribution for each animal taken across cross-validation splits/sessions, ± the SEM across sessions.

We performed both static and time-varying versions of this analysis. In the time-varying analysis, which describes the temporal evolution of category information in eye movements, classifiers were fit on eye positions at each single time-point. In the static analysis, classifiers were fit on the flattened vector of eye positions measured during the delay period. In both cases, eye position vectors were temporally down-sampled (resolution = 50 ms) before fitting classifiers or assessing their performance.

Two kinds of control analysis were performed to place classification results in context. First, we performed a standard shuffle control: in each session, we repeatedly shuffled the category labels assigned to individual trials and performed the same analysis. Second, we performed a boundary-shift control: Instead of shuffling the labels assigned to individual trials, we instead assessed classifier performance using labels assigned according to a category boundary orthogonal to the true boundary. Unlike the standard shuffle control, this analysis preserves groupings of trials with the same motion direction, as well as the overall structure of the relationship between direction and category labels (the counterfactual orthogonal category boundary carves up the stimulus space in the same way as the true boundary, but with maximally different groupings of directions). If separation in eye movements reflects separation according to the presented direction of motion, rather than the learned category per se, then classification performance should be high under the boundary-shift control. If the separation specifically reflects the learned category identity of the stimuli, however, then classification performance should be low under the boundary-shift control.

Congruence index. To quantify how closely eye movements on error trials for each category aligned with those of correct trials of that category, we devised a congruence index (CI). For each category x, we computed the ratio $(ExCx' - ExCx)/(ExCx' + ExCx)$, where $ExCx$ ($ExCx'$) is the distance of the average eye trace on error trials to the average eye trace on correct trials with the same (opposite) category. The index, the mean of this ratio taken across categories, ranges between −1 and 1. If eye positions on error trials are identical to those on correct trials of the *same* category, the index takes a value of 1; if eye positions on error trials are identical to those on correct trials of the *opposite* category, the index takes a value of −1; if eye positions on error trials are equidistant to those on correct trials of same and opposite categories, the index takes a value of 0. Eye movement data were assembled and combined across sessions as above (*Quantifying Separation in Eye Position by Category*). To assess how CI computed on error trials compared to that arising from variability in eye movements on correct trials, we performed the same analysis on a held-out subsample of correct trials.

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

Identifying shift-coding and category-coding directions in population activity space. We adopted an approach inspired by ref. 43 to identify coding directions in population activity space for gaze shifts (**sd**) and categories (**cd**) (Fig. 5). We defined the category-coding direction **cd** as the N-dimensional vector $\mathbf{r}_{c0} - \mathbf{r}_{c1}$, the difference between the population response in the 200 ms before eye movements began to reflect categories for trials of category 0 vs. category 1. Similarly, we defined **sd** as the single direction in N-dimensional activity capturing maximal variation related to where the eyes will move–the first principal component of activity in the 150 ms before eye movements in different directions made during MGS, where animals made large eye movements (~8 dva) in eight unique directions.

Angles between **sd** and **cd** were computed using a variant of the standard formula $\theta = \arccos(\mathbf{sd} \cdot \mathbf{cd})/(|\mathbf{sd}||\mathbf{cd}|)$. Because the orientation of **sd** and **cd** reflects an arbitrary choice of ordering in computing the difference (e.g., $\mathbf{r}_{c0} - \mathbf{r}_{c1}$ rather than $\mathbf{r}_{c1} - \mathbf{r}_{c0}$), the sign of the dot product in the numerator is not meaningful. We thus instead report $\theta = \arccos(|\mathbf{sd} \cdot \mathbf{cd}|)/(|\mathbf{sd}||\mathbf{cd}|)$–the acute angle between **sd** and **cd**, which is insensitive to the ordering of conditions when computing the difference vectors. θ ranges between 0° and 90°, with $\theta = 0$ degrees indicating that **sd** and **cd** point in exactly the same direction and $\theta = 90°$ indicating that **sd** and **cd** are orthogonal. To determine when a given value of θ indicates that **sd** and **cd** are more or less orthogonal than expected by chance, we also estimated θ_{shuf}, the distribution of θ under the null hypothesis that **sd** and **cd** are randomly oriented with respect to one another. We obtained θ_{shuf} by computing the angle between **sd** and a distribution of $\mathbf{cd}_{rand}$, obtained using the same procedure as **cd** but with respect to randomly shuffled category labels. Unless specified otherwise, significance of the difference between θ and θ_{shuf} was assessed via a one-sided bootstrap test.

We performed the same analysis–compute **sd** and **cd**, and the angle between them–across 1,000 independently generated pseudopopulations, constructed using the procedure described above in *Pseudopopulation Decoding of Gaze Shift Direction* (N pseudotrials per sample stimulus/gaze shift direction = 20). In particular, we 1) estimated **sd** in each of 1,000 pseudopopulations; 2) estimated **cd** in each of 1,000 additional pseudopopulations, shuffling the true category labels each time independently at random to also obtain $\mathbf{cd}_{rand}$ for that pseudopopulation; 3) computed the angle between each **sd** and $\mathbf{cd}/\mathbf{cd}_{rand}$. This yielded two distributions of angles–1,000 angles between **sd** and **cd**, and 1,000 angles between **sd** and $\mathbf{cd}_{rand}$. Error bars in the scatterplot shown in Fig. 5*C* reflect the SD of the 1,000 angles with respect to $\mathbf{cd}/\mathbf{cd}_{rand}$.

Quantifying impacts of muscimol inactivation on category-correlated eye movements. To quantify the impact of SC/LIP inactivation on category-correlated eye movements, we performed a simplified version of the analysis described in *Quantifying Separation in Eye Positions by Category*. First, we performed the analysis separately for behavior in each inactivation and control (sham or saline) session, and report variability (SEM) across sessions in each group (N = # sessions). For inactivation sessions, we only used trials corresponding to the "treatment" condition: during SC inactivations, only including trials performed after muscimol infusion, and for LIP inactivations, only including trials where the sample stimulus was positioned in the hemifield contralateral to the inactivated hemisphere. Second, to compute BCD and WCD, we split trials in each session by category rather than by direction to yield four traces of average difference in eye position: category 0 trials vs. held-out category 0 trials, category 0 trials vs. held-out category 1 trials, category 1 trials vs. held-out category 0 trials, and category 1 trials vs. held-out category 1 trials. Trials in each category were split in half at random (without replacement) to generate held-out trial sets for each session. This process was repeated 1,000 times. To capture the effect of inactivation on separation in eye movements using a single scalar value, we computed the peak difference D_i between BCD and WCD during the delay period in each control/inactivation session *i*. Our measure of effect size was the ratio $(D_{inactivation} - D_{control})/D_{control}$–the amount by which eye movement separation according to category changed during inactivation relative to control sessions, normalized by the level of separation during control sessions to yield a percentage change. A reduction in separation in eye movements by category corresponds to a negatively signed effect size. Significance of the difference between control and inactivation was assessed using a two-sided Wilcoxon signed-rank test on $D_{control}$ vs. $D_{inactivation}$, with $N_{control}$ and $N_{inactivation}$ equal to the number of sessions of control behavior and muscimol infusion respectively.

Determining features of eye movements (microsaccades and fixational drift) that differ between categories. To determine which aspects of fixational eye movements contribute to the category-specific gaze patterns we observed, we broke each correctly completed DMC trial's eye movements into microsaccades (identified using a velocity-based procedure modified from refs. 4, 39, and 40, and verified by extensive manual inspection) and periods of fixational drift (stretches between microsaccades without any ballistic eye movements). We also labeled each microsaccade and drift with its key descriptive features (amplitude and direction for each microsaccade; speed and direction for each drift). We then assessed the significance of the difference between categories for each of these features within each session through time.

Concretely, this entailed: 1) selecting all microsaccades and gaze shifts within a specified time window *w*; 2) splitting these eye movements according to the sample stimulus category from the trial on which they occurred, yielding two feature distributions for each of the four features, 1 per category; 3) testing the significance of the difference in location/mean between categories for each feature. For this analysis, we used windows of length 150 ms, stepped forward by 25 ms. To ensure consistency across datasets despite minor variation in task event durations/timings, we only used the 1,500 ms extending from 600 before to 900 ms after sample stimulus offset, a period that captured category-specific gaze patterns across all datasets. We tested for significance of the difference in direction for microsaccades/drifts using a permutation test based on Watson's nonparametric U^2 test for circular data. We used the analogous test for independent linear quantities (the Mann–Whitney U test) to test for significant differences in microsaccade amplitude and drift speed between categories. We evaluated the significance of the difference in microsaccade rate between categories using a shuffling procedure: Within every session, we shuffled the category labels assigned to each trial, for each shuffle recomputing the difference in microsaccade rate through time. Using the resulting distribution of rate difference time-courses, we computed a *P*-value at each timepoint (% of the shuffled samples with as large of a difference as the true data). As elsewhere, all time-courses of *P*-values were adjusted using the Benjamini–Hochberg procedure to control the false discovery rate.

Data, Materials, and Software Availability. Data analyzed in this study will be posted to FigShare (https://www.figshare.com) at the time of publication, along with analysis code for generating figures (44).

ACKNOWLEDGMENTS. We thank Drs. W. Jeffrey Johnston, Krithika Mohan, and Vinay Shirhatti for feedback on an earlier version of this manuscript. This work was supported by NIH R01EY019041 and DOD Vannevar Bush Faculty Fellowship N000141912001 (D.J.F.), as well as support from the Margot and Thomas Pritzker Family Foundation. We also thank previous members of the Freedman and Assad laboratories for access to their data for analysis, especially Drs. Jamie Fitzgerald, Sruthi Swaminathan, Chris Rishel, Arup Sarma, Krithika Mohan, Yang Zhou, and Barbara Peysakhovich.

Author affiliations: [a]Department of Neurobiology, The University of Chicago, Chicago, IL 60637; and [b]Neuroscience Institute, The University of Chicago, Chicago, IL 60637

1. D. J. Freedman, J. A. Assad, Experience-dependent representation of visual categories in parietal cortex. *Nature* **443**, 85–88 (2006).
2. Y. Zhou, K. Mohan, D. J. Freedman, Abstract encoding of categorical decisions in medial superior temporal and lateral intraparietal cortices. *J. Neurosci.* **42**, 9069–9081 (2022).
3. S. K. Swaminathan, D. J. Freedman, Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. *Nat. Neurosci.* **15**, 315–320 (2012).
4. Y. Zhou, O. Zhu, D. J. Freedman, Posterior parietal cortex plays a causal role in abstract memory-based visual categorical decisions. *J. Neurosci.* **43**, 4315–4328 (2023).
5. B. Peysakhovich *et al.*, Primate superior colliculus is causally engaged in abstract higher-order cognition. *Nat. Neurosci.* **27**, 1999–2008 (2024), 10.1038/s41593-024-01744-x.
6. R. Engbert, R. Kliegl, Microsaccades uncover the orientation of covert attention. *Vis. Res.* **43**, 1035–1045 (2003).
7. S. Martinez-Conde, J. Otero-Millan, S. L. Macknik, The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nat. Rev. Neurosci.* **14**, 83–96 (2013).
8. A. S. Barnhart, F. M. Costela, S. Martinez-Conde, S. L. Macknik, S. D. Goldinger, Microsaccades reflect the dynamics of misdirected attention in magic. *J. Eye Mov. Res.* **12**, 7 (2019).

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.

9. L. L. Di Stasi *et al.*, Microsaccade and drift dynamics reflect mental fatigue. *Eur. J. Neurosci.* **38**, 2389–2398 (2013).
10. L. L. Di Stasi *et al.*, Saccadic eye movement metrics reflect surgical residents' fatigue. *Ann. Surg.* **259**, 824–829 (2014).
11. L. L. Di Stasi *et al.*, Effects of long and short simulated flights on the saccadic eye movement velocity of aviators. *Physiol. Behav.* **153**, 91–96 (2016).
12. P. Mostert *et al.*, Eye movement-related confounds in neural decoding of visual working memory representations. *eNeuro* **5**, ENEURO.0401-17.2018 (2018).
13. S. C. Quax, N. Dijkstra, M. J. van Staveren, S. E. Bosch, M. A. J. van Gerven, Eye movements explain decodability during perception and cued attention in MEG. *NeuroImage* **195**, 444–453 (2019).
14. N. Zokaei, A. G. Board, S. G. Manohar, A. C. Nobre, Modulation of the pupillary response by the content of visual working memory. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22802–22810 (2019).
15. F. van Ede, S. R. Chekroud, A. C. Nobre, Human gaze tracks attentional focusing in memorized visual space. *Nat. Hum. Behav.* **3**, 462–470 (2019).
16. J. Linde-Domingo, B. Spitzer, Geometry of visuospatial working memory information in miniature gaze patterns. *Nat. Hum. Behav.* **8**, 336–348 (2024).
17. A. Libby, T. J. Buschman, Rotational dynamics reduce interference between sensory and memory representations. *Nat. Neurosci.* **24**, 715–726 (2021).
18. D. J. Freedman, M. Riesenhuber, T. Poggio, E. K. Miller, Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312–316 (2001).
19. V. B. Mountcastle, J. C. Lynch, A. Georgopoulos, H. Sakata, C. Acuna, Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. *J. Neurophysiol.* **38**, 871–908 (1975).
20. D. L. Robinson, M. E. Goldberg, G. B. Stanton, Parietal association cortex in the primate: Sensory mechanisms and behavioral modulations. *J. Neurophysiol.* **41**, 910–932 (1978).
21. R. A. Andersen, G. K. Essick, R. M. Siegel, Neurons of area 7 activated by both visual stimuli and oculomotor behavior. *Exp. Brain Res.* **67**, 316–322 (1987).
22. J. K. Fitzgerald, D. J. Freedman, J. A. Assad, Generalized associative representations in parietal cortex. *Nat. Neurosci.* **14**, 1075–1079 (2011).
23. C. A. Rishel, G. Huang, D. J. Freedman, Independent category and spatial encoding in parietal cortex. *Neuron* **77**, 969–979 (2013).
24. A. Sarma, N. Y. Masse, X.-J. Wang, D. J. Freedman, Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices. *Nat. Neurosci.* **19**, 143–149 (2016).
25. K. Mohan, O. Zhu, D. J. Freedman, Interaction between neuronal encoding and population dynamics during categorization task switching in parietal cortex. *Neuron* **109**, 700–712.e4 (2021).
26. W. Bair, L. P. O'Keefe, The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Vis. Neurosci.* **15**, 779–786 (1998).
27. D. J. Freedman, M. Riesenhuber, T. Poggio, E. K. Miller, A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci.* **23**, 5235–5246 (2003).
28. S. K. Swaminathan, N. Y. Masse, D. J. Freedman, A comparison of lateral and medial intraparietal areas during a visual categorization task. *J. Neurosci.* **33**, 13157–13170 (2013).
29. B. C. Talluri *et al.*, Activity in primate visual cortex is minimally driven by spontaneous movements. *Nat. Neurosci.* **26**, 1953–1959 (2023).
30. Z. M. Hafed, L. Goffart, R. J. Krauzlis, A neural mechanism for microsaccade generation in the primate superior colliculus. *Science* **323**, 940–943 (2009).
31. N. Y. Masse, J. M. Hodnefield, D. J. Freedman, Mnemonic encoding and cortical organization in parietal and prefrontal cortices. *J. Neurosci.* **37**, 6098–6112 (2017).
32. K. W. Latimer, D. J. Freedman, Low-dimensional encoding of decisions in parietal cortex reflects long-term training history. *Nat. Commun.* **14**, 1010 (2023).
33. S. Musall, M. T. Kaufman, A. L. Juavinett, S. Gluf, A. K. Churchland, Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* **22**, 1677–1686 (2019).
34. C. Stringer *et al.*, Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, eaav7893 (2019).
35. S. Tremblay, C. Testard, R. W. DiTullio, J. Inchauspé, M. Petrides, Neural cognitive signals during spontaneous movements in the macaque. *Nat. Neurosci.* **26**, 295–305 (2023).
36. D. A. Leopold, N. K. Logothetis, Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp. Brain Res.* **123**, 341–345 (1998).
37. S. Martinez-Conde, S. L. Macknik, D. H. Hubel, Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nat. Neurosci.* **3**, 251–258 (2000).
38. D. M. Snodderly, I. Kagan, M. Gur, Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Vis. Neurosci.* **18**, 259–277 (2001).
39. S. Martinez-Conde, S. L. Macknik, D. H. Hubel, The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13920–13925 (2002).
40. T. M. Herrington *et al.*, The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. *J. Neurosci.* **29**, 5793–5805 (2009).
41. G. Mongillo, O. Barak, M. Tsodyks, Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
42. S. Ben Hamed, J.-R. Duhamel, F. Bremmer, W. Graf, Representation of the visual field in the lateral intraparietal area of macaque monkeys: A quantitative receptive field analysis. *Exp. Brain Res.* **140**, 127–144 (2001).
43. N. Li, K. Daie, K. Svoboda, S. Druckmann, Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
44. M. C. Rosen, D. J. Freedman, Data from "Shared files for Multiplexing of Cognitive Encoding by Oculomotor Networks Leads to Incidental Gaze Shifts." Figshare. https://dx.doi.org/10.6084/m9.figshare.28651907. Deposited 24 March 2025.

Downloaded from https://www.pnas.org by UNIVERSITY OF CHICAGO; THE JOHN CRERAR LIBRARY on April 8, 2025 from IP address 128.135.93.238.