

# Linking Objects to Actions: Encoding of Target Object and Grasping Strategy in Primate Ventral Premotor Cortex

Carlos E. Vargas-Irwin,<sup>1</sup> Lachlan Franquemont,<sup>1</sup> Michael J. Black,<sup>2</sup> and John P. Donoghue<sup>1</sup>

<sup>1</sup>Department of Neuroscience, Brown University, Providence, Rhode Island 02912, and <sup>2</sup>Max Planck Institute for Intelligent Systems, Tübingen, 72076, Germany

Neural activity in ventral premotor cortex (PMv) has been associated with the process of matching perceived objects with the motor commands needed to grasp them. It remains unclear how PMv networks can flexibly link percepts of objects affording multiple grasp options into a final desired hand action. Here, we use a relational encoding approach to track the functional state of PMv neuronal ensembles in macaque monkeys through the process of passive viewing, grip planning, and grasping movement execution. We used objects affording multiple possible grip strategies. The task included separate instructed delay periods for object presentation and grip instruction. This approach allowed us to distinguish responses elicited by the visual presentation of the objects from those associated with selecting a given motor plan for grasping. We show that PMv continuously incorporates information related to object shape and grip strategy as it becomes available, revealing a transition from a set of ensemble states initially most closely related to objects, to a new set of ensemble patterns reflecting unique object-grip combinations. These results suggest that PMv dynamically combines percepts, gradually navigating toward activity patterns associated with specific volitional actions, rather than directly mapping perceptual object properties onto categorical grip representations. Our results support the idea that PMv is part of a network that dynamically computes motor plans from perceptual information.

**Key words:** grasping; macaque; microelectrode array; SSIMS; ventral premotor

## Significance Statement

The present work demonstrates that the activity of groups of neurons in primate ventral premotor cortex reflects information related to visually presented objects, as well as the motor strategy used to grasp them, linking individual objects to multiple possible grips. PMv could provide useful control signals for neuroprosthetic assistive devices designed to interact with objects in a flexible way.

## Introduction

The ability to interact with a large variety of real-world objects in flexible and creative ways is one of the most salient traits of the primate cerebral motor system, and has arguably played a pivotal role in the evolutionary success of our species (Napier, 1980; Lemon, 1993). Upon looking at an object, we can quickly plan and execute a variety of movements to grasp and manipulate it in

a context-appropriate manner. The ventral premotor cortex (PMv) is a key node in the parietofrontal network specialized in transforming visual information representing the shape of objects into hand postures best suited for grasping (Fagg and Arbib, 1998; Luppino et al., 1999; Rizzolatti and Luppino, 2001; Tanné-Gariépy et al., 2002; Brochier and Umiltà, 2007; Theys et al., 2012; 2013). Transient PMv inactivation impairs hand reshaping preceding grasping, but does not impair reaching (Fogassi et al., 2001). Further, individual PMv neurons display grip-specific activation during grasping; differential responses to various objects are also observed during instructed delays preceding grasping movements, and even during visual object fixation when no movement is intended (Godschalk et al., 1981; Murata et al., 1997; Raos et al., 2006; Umiltà et al., 2007; Fluet et al., 2010; Carpaneto et al., 2011; Townsend et al., 2011; Schaffelhofer et al., 2015).

Although most real-world objects can be grasped in many different ways, previous experimental work has focused on examining single-neuron activity during actions in which each target

Received April 22, 2015; revised June 26, 2015; accepted July 1, 2015.

Author contributions: C.E.V.-I., L.F., M.J.B., and J.P.D. designed research; C.E.V.-I. and L.F. performed research; C.E.V.-I. contributed unpublished reagents/analytic tools; C.E.V.-I. and L.F. analyzed data; C.E.V.-I., M.J.B., and J.P.D. wrote the paper.

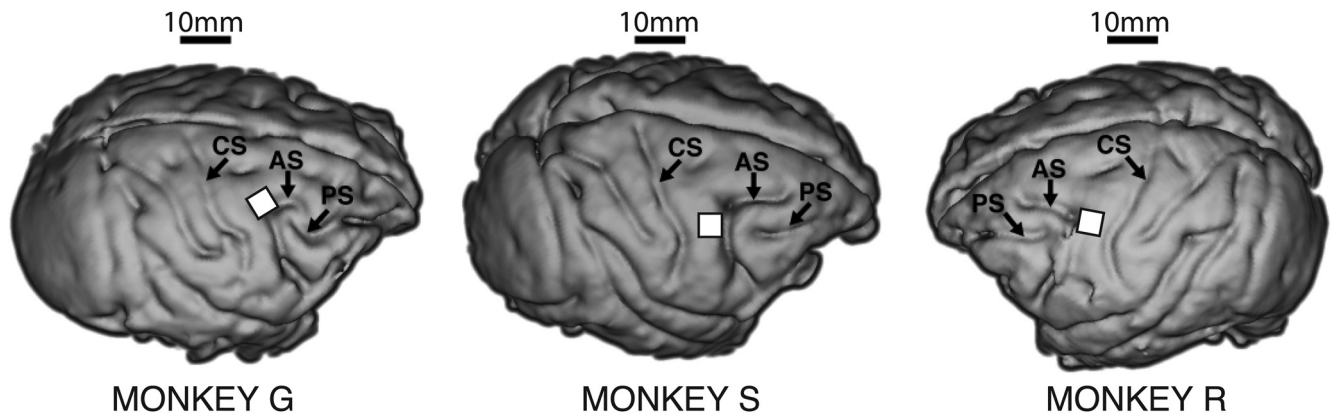
This work was supported by VA-Rehab R&D, NINDS-Javits (NS25074), DARPA REPAIR (N66001-10-C-2010), and the Katie Samson Foundation. We thank Corey Triebwasser, Michelle Nevor, John Murphy, and Allan Rydberg for their assistance with animal care and instrumentation design.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr Carlos E. Vargas-Irwin, Department of Neuroscience, Sidney E. Frank Hall for Life Science, 185 Meeting Street, Brown University, Providence, RI 02912. E-mail: Carlos\_Vargas\_Irwin@Brown.edu.

DOI:10.1523/JNEUROSCI.1574-15.2015

Copyright © 2015 the authors 0270-6474/15/3510888-10\$15.00/0



**Figure 1.** Multielectrode array implantation sites for the three NHPs used in the study. A  $10 \times 10$  microelectrode array ( $400 \mu\text{m}$  spacing) was chronically implanted at each site (Blackrock Microsystems). In all cases, arrays were positioned just anterior to the genu of the arcuate sulcus at the (mediolateral) level of the principal sulcus. 3D rendering of the cortical surfaces was generated based on 3T MRI scans (Brown MRI Research Facility) using BrainSight software (Rouge Research). CS, Central sulcus; AS, arcuate sulcus; PS, principal sulcus.

object was associated with a single well-practiced grip strategy. This has promoted the view that objects elicit PMv activity specifying a particular grip strategy, suggesting that the computation linking objects to actions requires a mapping of percepts to sets of cells representing specific grip categories. However, in the natural world objects can be grasped in many different ways depending on the situation. It is not clear how PMv circuits could flexibly link either the same action to multiple objects or multiple actions to the same object. In this study, we used multielectrode arrays to record neural ensemble activity in PMv of nonhuman primates (NHPs) engaged in a grasp and lift task where each of two target objects could be gripped in two different ways. Objects shared one grip type (power) and had one unique grip each (precision or key). The task included a sequence of two instructed delay periods: during the first, an object was viewed without knowledge of the correct grip type, during the second, colored lights were used to cue the correct grip. This approach allowed us to distinguish responses elicited by the visual presentation of the objects from those associated with selecting a way to grasp it. We use spike train similarity (SSIMS) analysis (Vargas-Irwin et al., 2015) to track the progression of PMv ensemble states leading from object viewing to grasping. This method allows us to quantify the similarity of the ensemble activity patterns across trials including those elicited by grasping a given object in different ways, as well as those associated with grasping different objects using the same grip.

Our results show that PMv ensemble activity dynamically evolves throughout the process of selecting and executing a grasping action. Our task design elicits a transition from a set of ensemble states most closely associated with the target objects to a separate set of states displaying object and grip information, reflecting the neural computations linking objects to actions.

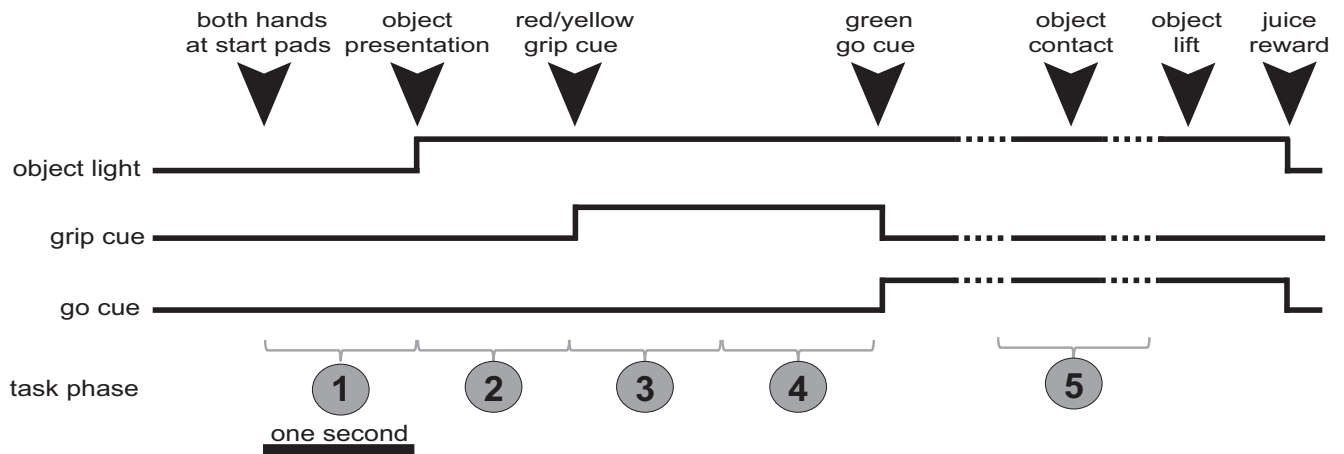
## Materials and Methods

**Neural population activity recording.** Neural ensemble activity was recorded from three male macaque monkeys (9–12 kg) using chronically implanted microelectrode arrays (Blackrock Microsystems). Animals were maintained in a facility approved by the Association for Assessment and Accreditation of Laboratory Animal Care, National Institutes of Health (AAALAC, NIH). All surgery was performed using standard sterile procedures in an approved animal surgical facility. The microelectrode arrays were implanted into the functionally defined PMv upper limb area accessible on the cortical surface (just posterior to the genu of the arcuate sulcus, at the level of the principal sulcus); this region on the periarculate convexity, corresponding to the F4–F5c boundary (using the

terminology of Belmalih et al., 2009), has been reliably found to contain neurons related to distal upper limb movements (Kurata and Tanji, 1986; Rizzolatti et al., 1988; Raos et al., 2006; Umiltà et al., 2007; Nelissen and Vanduffel, 2011). Each electrode array contained a  $4 \times 4$  mm grid with 96 active silicon-based electrodes 1 mm in length and spaced  $400 \mu\text{m}$  apart (for details on array structure and surgical procedures, see Suner et al., 2005; Barrese et al., 2013). The location of each PMv array on the cortex is displayed in Figure 1. Additional arrays were implanted in primary motor cortex and/or dorsal premotor cortex (data from these arrays not included in the analysis presented). Monkeys were head fixed during recording. Data acquisition and storage were performed using a Cerebus multichannel data acquisition system (Blackrock Microsystems). Differences in spike waveform shape were used to identify single-unit activity using custom-made software using template matching with spike overlap resolution (Vargas-Irwin and Donoghue, 2007). Only units exceeding a signal-to-noise ratio of 1.2 were included in the analysis ( $\text{SNR} = \text{mean spike amplitude}/95\%$  confidence interval for voltage during nonspiking periods).

**Behavioral task.** We recorded the activity of PMv neuronal ensembles using chronically implanted microelectrode arrays in three NHPs (rhesus macaques) trained to perform a cued grasping task with instructed delay (CGID). Recording sites are shown in Figure 1. Four recording sessions were performed for each NHP. In each session, between 49 and 135 neurons were simultaneously recorded in PMv.

The temporal sequence of the task is shown in Figure 2. To initiate a trial, the head-fixed monkeys, sitting in a primate chair in a fully darkened room, held their hands on two touch-sensitive pads for one second. One of two possible objects was presented in front of them using a custom made turntable device and then illuminated by a spotlight positioned above. After a further one second delay, one of two grip-type instructions was provided using colored lights. Monkeys were trained to grasp objects in one of two ways using the following conditional associations: a red light instructed a power grip (common to both objects), and a yellow light instructed either a key or precision grip (depending on which object was present). The monkeys were required to keep their hands in contact with start pads until a “go” cue (green light) was delivered (which happened 2 s after the grip instruction cue). They then reached forward, grasped the object with the instructed grip and lifted it. The objects were mounted on sliding rails that allowed 2 cm of vertical displacement; a fruit juice reward was given if the objects were held at the top position for  $>200$  ms. To detect whether the correct grip was being used, each object was instrumented with four capacitive sensors (Qtouch sensors, Atmel). For all grips, two points of contact on opposite sides of the object were required (with an independent pair of sensors for each grip). Both objects included an identical 10-cm-long, 3.5-cm-diameter cylinder. For Object A, the cylinder was joined to a horizontal disk with a diameter of 7.5 cm and thickness of 0.5 cm. In Object B, the cylinder was joined to a vertical plate measuring  $10 \times 7 \times 0.5$  cm. In Monkey G, the



**Figure 2.** CGID task. Task timeline, showing the onset and offset of the object light, grip cue, and go cue. There are four types of trials, representing combinations of two objects and three grips (with both objects sharing one grip in common). For analysis purposes, the task was divided into five 1 s time periods.

cylinder was the top part of the object (Fig. 3). In subsequent experiments with Monkeys S and R the cylinder was positioned at the bottom (we found that this configuration was easier for the monkeys to visualize and grasp while in head fixation). For each recording session, the monkeys performed between 80 and 150 trials, with between 15 and 40 exemplars of each grip/object combination. Objects and grips were pseudorandomly interleaved throughout the session (keeping trial numbers for all combinations approximately consistent). In summary, for each trial the monkeys were presented with one of the two objects and, depending on subsequent instructional cues, grasped them either in a similar way (power grip) or a different way (precision or key grip).

**SSIMS analysis.** We used spike train similarity space analysis (SSIMS; as described by Vargas-Irwin et al., 2015) to quantify object and grip-related effects on neural activity at each of the task phases. SSIMS is an example of relational data analysis that uses estimates of similarity to describe the relationship between multiple data samples (Lehky et al., 2013). Note that mapping neural data into a SSIMS projection is done without any information regarding behavioral condition (unsupervised dimensionality reduction). Instead, the method relies on the intrinsic properties of the neural data and does not require an explicit model of the relationship between neural activity and external variables. The SSIMS algorithm begins by embedding the neural data into a high-dimensional pairwise similarity space, and then projects the data into a more compact (low-dimensional) representation, which facilitates statistical analysis as well as data visualization while still capturing the relationship between individual data points. A neural activity pattern is represented by a point in the resulting SSIMS projection (which can be defined over one or many neurons simultaneously). The distance between points denotes the degree of similarity between the activity patterns they represent. Two identical firing patterns correspond to the same point in this space; the more different they are, the farther apart they lie in the SSIMS projection.

The following operations are performed to generate a SSIMS projection of dimensionality  $d$  representing the activity of  $n$  neurons recorded over  $t$  trials: (1) generate a high-dimensional embedding for the data using the pairwise distances between the spike trains recorded for each neuron. We use the spike train metric proposed by Victor and Purpura (1996) to evaluate similarity. This method evaluates differences in spike trains based on how spikes in one train have to be shifted, deleted, or inserted to make it precisely match another train (similar to the “edit distance” or “Levenshtein distance” used to compare letter strings or genetic sequences). The cost function used was adjusted such that the cost of shifting a spike by  $>100$  ms was equivalent to inserting and deleting a spike ( $1/q = 100$  ms). The activity pattern for one neuron during a single trial is represented by a similarity vector of length  $t$ , representing the distance to each of the other trials. When performing ensemble-level analysis, vectors from multiple neurons are concatenated, resulting in a set of  $t$  ensemble similarity vectors with dimensionality  $n \cdot t$ . We define the state of the ensemble as the combination of all the

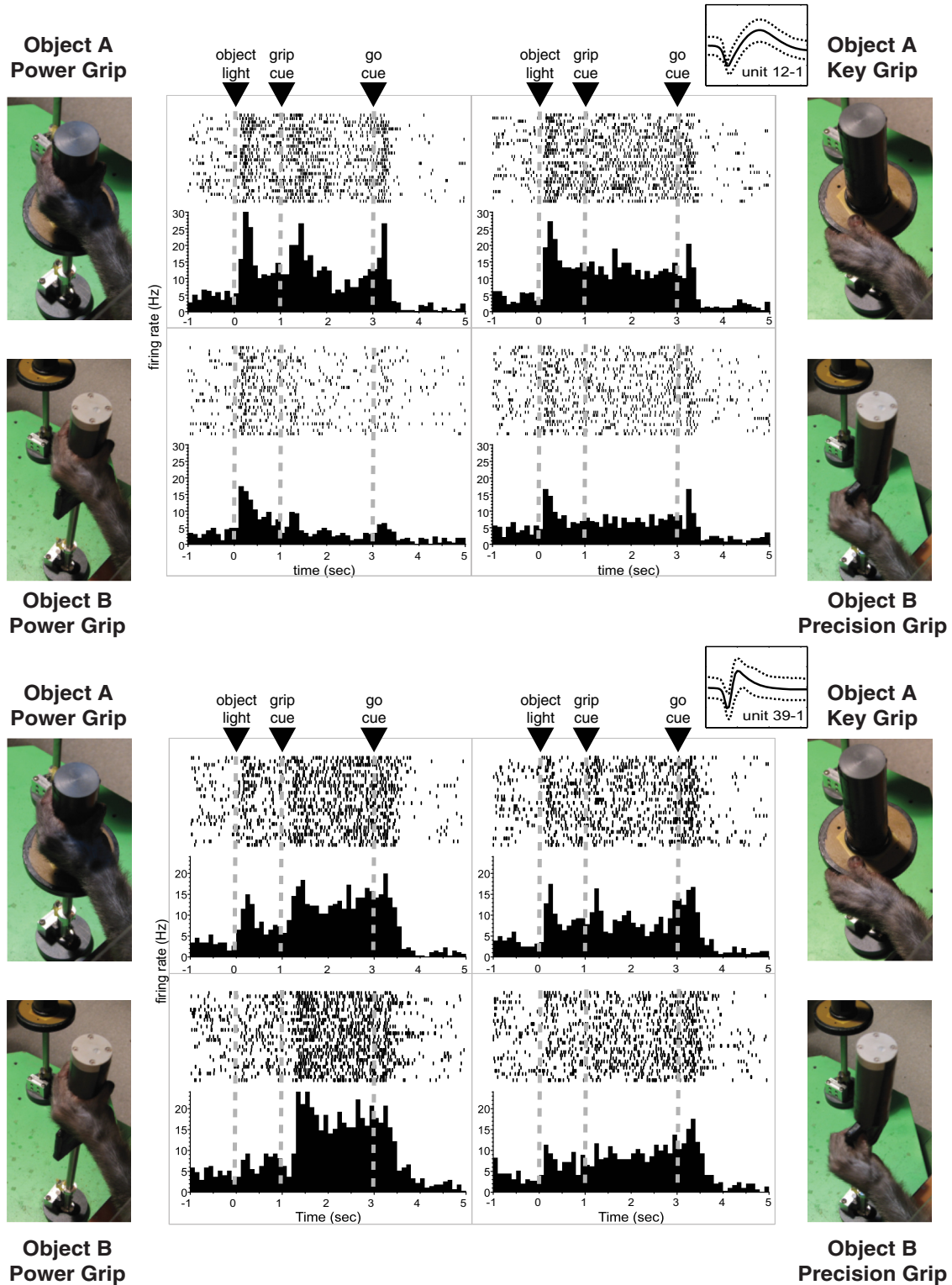
recorded spikes fired by all well isolated neurons during each of the 1 s time periods analyzed; two ensemble states are considered identical if all the spikes for each neuron occurred at exactly the same relative times. In the SSIMS representation, identical states are mapped to exactly the same point.

(2) Perform dimensionality reduction using the t-SNE algorithm developed by van der Maaten and Hinton (2008). The goal is to project the original high-dimensional data into a low-dimensional space of size  $d$  to facilitate visualization and analysis. The t-SNE algorithm uses explicit models of the “local neighborhoods” around each high-dimensional point to perform dimensionality reduction while preserving the relational structure of the data. The use of dynamic neighborhood sizes allows the algorithm to preserve local structure through areas of varying data density. The data are initially transformed to an intermediate number of dimensions using principal component analysis (for computational efficiency). In our implementation, we used a 50-dimensional space. Gradient descent is then used to minimize the KL-divergence between local neighborhood probability functions in the 50-dimensional PCA space and the target low dimensional SSIMS space (initialized using a second  $d$ -dimensional PCA projection). Our implementation of t-SNE used a perplexity setting of 30 (equivalent to the effective number of neighbors used to determine the Gaussian kernel used for density estimation).

Our previous work has demonstrated that the combination of spike metrics and t-SNE outperforms other alternatives, including similarity estimates based on binned spike counts and dimensionality reduction using PCA or multidimensional scaling (Vargas-Irwin et al., 2015). We have also shown that the SSIMS algorithm is robust over a wide range of parameter settings and capable of accurately classifying reaching and grasping movements from primate cortical recordings. MATLAB code for implementing the SSIMS algorithm is available upon request.

The present work uses SSIMS projections in four different ways:

(1) Evaluating single-neuron properties: the (Euclidian) distance between low-dimensional SSIMS projections can be used to estimate the similarity between neural activity patterns. If neural activity patterns differ between two experimental conditions, trials within the same condition will cluster closer together, whereas trials in different conditions will be further apart. This phenomenon can be quantified by comparing the distribution of within- and between-category separation in the SSIMS projection. Units for which the median distance for trials in the same category (i.e., object or grip) was significantly smaller than the median distance for trials in different categories were considered to have significant clustering (Kruskall–Wallis,  $p < 0.05$ ). Note that this approach goes beyond testing for differences in firing rate, because it also incorporates the timing of each spike, allowing us to evaluate relatively large time windows without binning. This technique was used to compare activity associated with the same object and different grips, or for the same grip and different objects. In the case of grip-related clustering,



**Figure 3.** Object and grip related information is reflected in the firing pattern of individual neurons. The activity patterns of two simultaneously recorded PMv neurons (Monkey G) are shown for each of the four object + grip combinations. In all cases, raster histograms are aligned to object presentation. Note that one neuron presents similar activity patterns for both power grips (bottom), whereas the other responds in a different way when the same grip is used with different target objects (top). Insets near the top right corner of each set of plots show the mean and 95% confidence interval of the waveforms assigned to each unit during spike sorting. The x-axis corresponds to 1.6 ms, whereas the y-axis represents  $\pm 100 \mu V$ . The SNR calculated for these units was 1.7 and 1.6 for units 12 and 39, respectively.

results from both objects were pooled together after adjusting the *p* values using Bonferroni correction for multiple comparisons. We used 15D SSIMS vectors ( $d = 15$ ) for statistical comparisons. Further increasing the dimensionality did not change the number of selective units, in agree-

ment with our previous decoding results using the SSIMS method (Vargas-Irwin et al., 2015).

(2) Graphically displaying the relationships (similarity) between single trials at a given point in time: SSIMS plots represent the activity of the



entire ensemble during each trial in an unbiased way, because the method does not take into account class labels. For visualization we use lower-dimensional projections of size  $d = 2$  or  $d = 3$ .

(3) Ensemble decoding: SSIMS projections can be easily adapted for discrete classification of neural activity patterns. Here, we use a simple nearest-neighbor classifier based on the Euclidean distances between SSIMS data points corresponding to individual trials. The classifiers were evaluated using leave-one-out cross validation. We used a 15D SSIMS space, which was sufficient to capture relevant data features (increasing dimensionality further did not improve decoding).

(4) Neural trajectories: it is also possible to use SSIMS to track the transition across different task phases, representing the temporal evolution of recorded ensemble activity through time (Churchland et al., 2007; note the overall goal is similar, although a different dimensionality reduction algorithm is used). Recall that the SSIMS algorithm represents a given ensemble activity pattern in terms of the similarity (distance metric) with other recorded patterns (a vector of similarity measurements). To generate neural trajectories, the activity patterns recorded over a sliding time window were compared with the full set of spike trains aligned to each task phase. Each time point is therefore represented by a vector with dimensionality  $n \cdot t \cdot p$ , where  $n$  is the number of neurons,  $t$  is the number of trials, and  $p$  is the number of task phases. Note that it is only necessary to run gradient descent on the set of spike trains at each task phase (which form the basis for the space), the resulting transform can then be applied to any number of  $n \cdot t \cdot p$  vectors representing other temporal alignments. In this manner, it is possible to evaluate the relationship between neural activity patterns at different time points in a single unified space.

## Results

### Single-unit responses in the CGID task

A total of 12 recording sessions were included in this study (4 sessions for each of the three NHP subjects). The number of well isolated single units identified per session ( $\text{SNR} > 1.2$ ) varied between 48 and 134, as detailed in Table 1 (mean = 84.6, SD = 34.9).

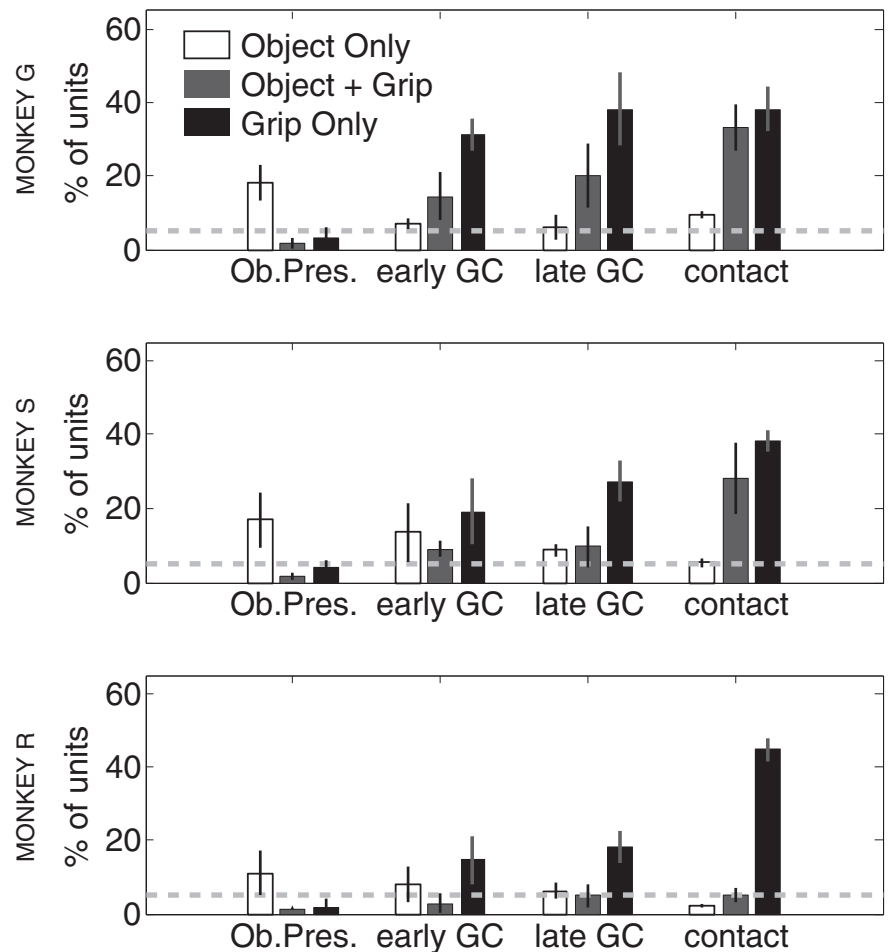
Single-unit activity in PMv displayed a wide variety of phasic and/or tonic responses associated with each of the task-relevant stimuli (Fig. 3). In addition to movement-related discharges, individual neurons typically displayed time-locked responses to the object presentation, as well as the grip instruction cues.

For analysis purposes we divided the task into five 1 s time periods (highlighted in Fig. 2): (1) intertrial interval (lights off, no cues), (2) object presentation phase (illumination), (3) early grip cue phase (first second after grip instruction, early planning), (4) late grip cue phase (late planning, following the early grip cue phase and immediately preceding the go cue), and (5) object contact phase (spanning 0.5 s before and after object contact was detected using capacitive sensors embedded in the object).

The first part of our analysis focused on the properties of individual neurons. To isolate the object and grip effects, we

**Table 1. Number of single units ( $\text{SNR} > 1.2$ ) identified in each session**

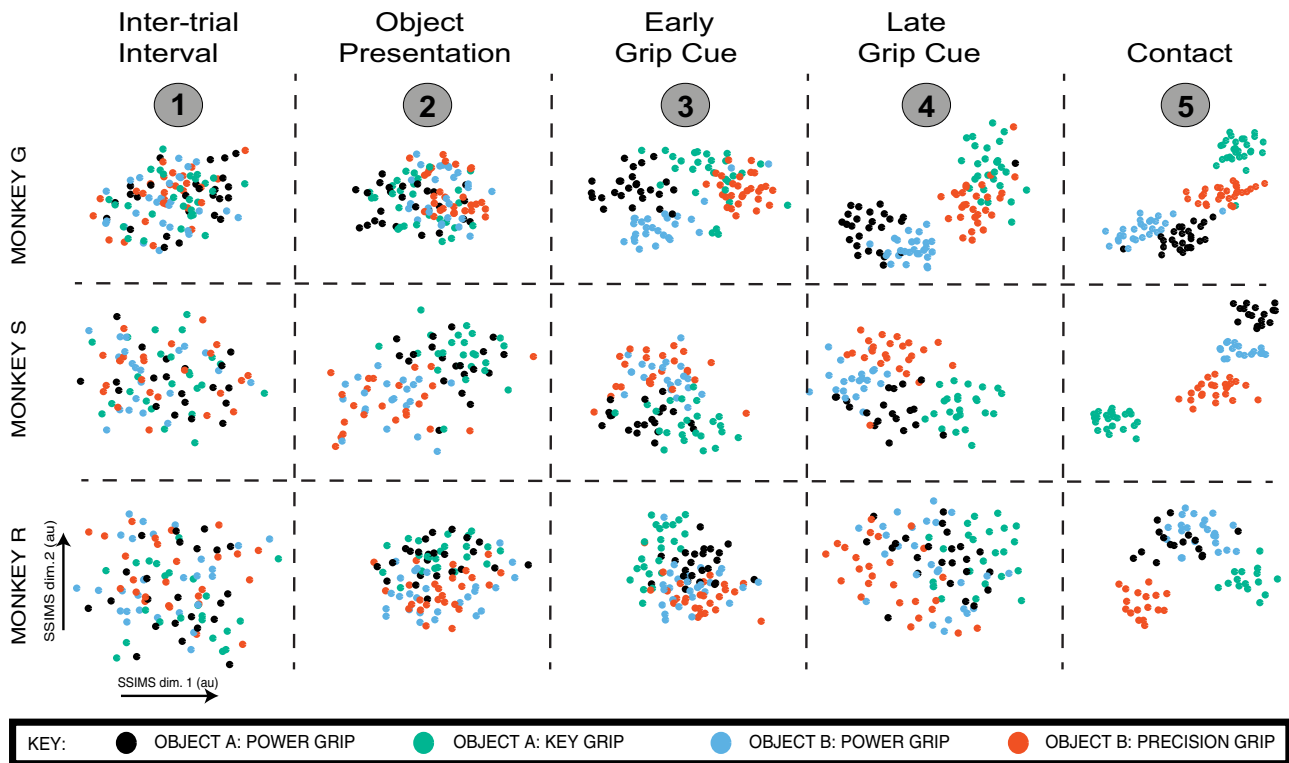
	Session 1	Session 2	Session 3	Session 4
Monkey G	60	49	50	48
Monkey R	62	81	72	76
Monkey S	116	138	129	134



**Figure 4.** Single-units displaying object- and grip-related information. SSIMS projections were used to determine whether individual neurons presented significant grip/object related information (significant difference in median within and between category SSIMS distances; KW,  $p < 0.05$ ). Object information was only evaluated for trials with the same grip (power) and different objects, whereas grip information was evaluated for different grips performed on the same object (key vs power or precision vs power). Each plot shows the number of neurons displaying exclusively object information, exclusively grip information, or both. The dashed line shows the number of false-positives expected for the given  $p$  value (5%). Data from each monkey is presented separately.

compared trials with the same grips performed on different objects (power grip trials), or trials with the same object grasped using different grips (power vs precision, or power vs key).

Significant clustering of SSIMS projections was evaluated in the manner described in the methods section to detect object or grip related information. The percentage of neurons displaying object and grip selectivity gradually increased through the different task stages, going from  $\sim 15\%$  to  $> 50\%$  (Fig. 4). During the object presentation period (before the grip had been instructed) only object-related information was observed above chance levels (representing  $\sim 15\%$  of all recorded neurons). Following the grip cue, the number of units showing grip related information far exceeded object-related neurons (Fig. 4). The number of exclusively object-related neurons decreased to near chance levels in the later stages of the task in all three monkeys. However, in two



**Figure 5.** Clustering of ensemble states according to object and grip strategy in the CGID task: single-session examples. Each panel clustering of ensemble states in SSIMS for a given task phase (arranged horizontally) and monkey (arranged vertically). The five task phases are shown in chronological order from left to right. There are four types of trials, representing combinations of two objects and three grips (with both objects sharing one grip in common). Colors denote type of trial, as noted in the key. Each point represents the ensemble state for a single trial, incorporating 1 s spike trains from each recorded neuron. The distance between the points denotes the relative similarity between the ensemble states. Note the progression from a random mix of trial types (colors), to clustering according to object, and then object/grip in the later stages of the task. Each SSIMS plot was generated independently to emphasize the relationship between individual trials during each task phase (axes are therefore not consistent between plots). Neural trajectories showing the transition between phases in a single unified space are shown in Figure 7.

monkeys (G and S) we observed a gradual increase in the number of neurons representing a combination of both object and grip information up to and including the action execution period. This trend was not observed in the third animal (R), where the number of neurons displaying object-related information dropped to chance levels by the object contact phase.

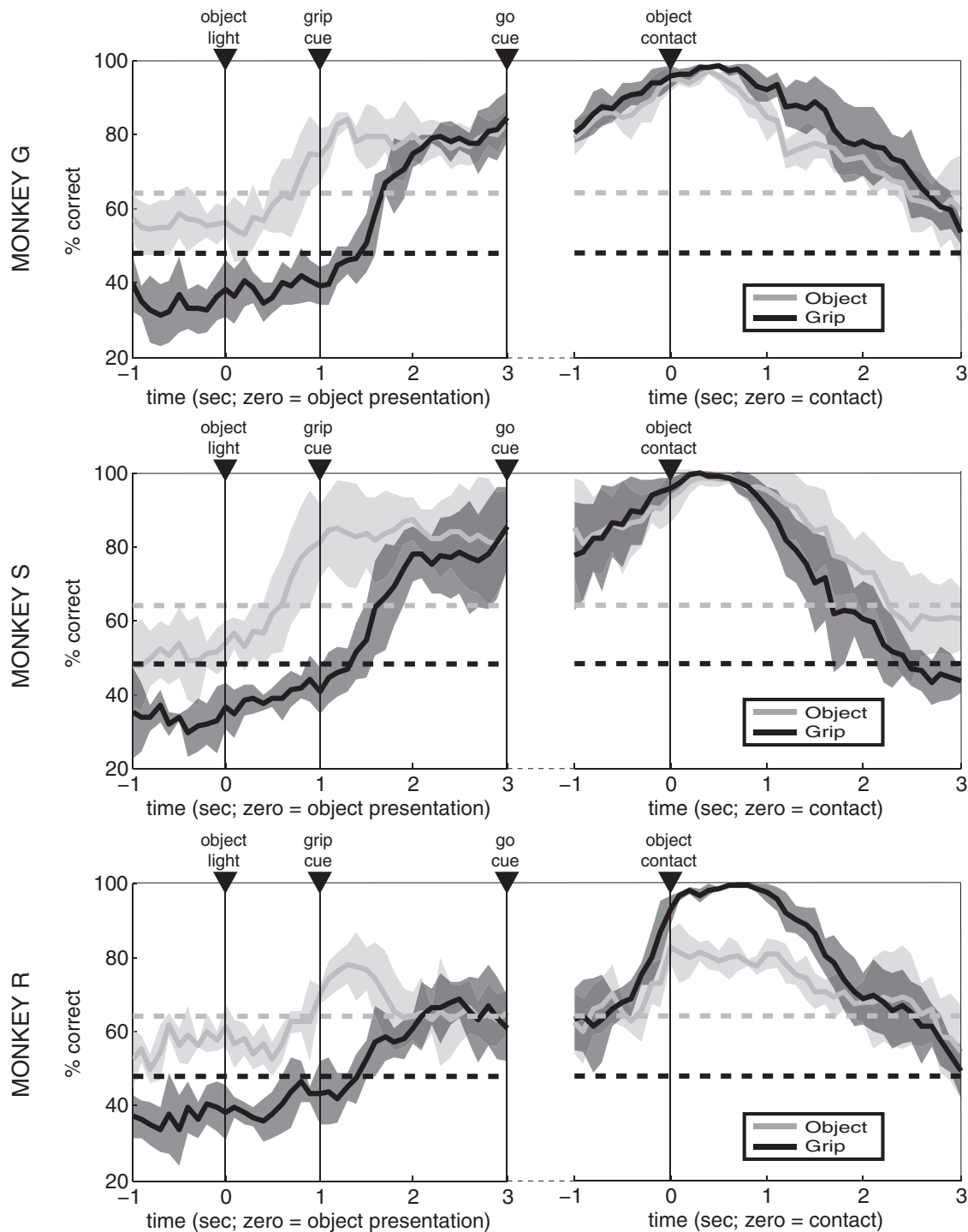
**Clustering neural ensemble states**

To display the informational content of PMv ensembles, we generated 2D SSIMS plots using the full set of simultaneously recorded neurons for each session (Fig. 5). Note that these plots are provided to facilitate visualization of the patterns present in the data, whereas the statistical analysis of clustering patterns is presented in subsequent figures. During phase 1 (the intertrial interval) the trial types were mixed randomly, and there was no clear clustering of similar activity patterns in the SSIMS projection, consistent with the lack of object or grip information available during this period. However, as more information was provided to the monkey, clustering of specific ensemble states emerged. During the object presentation period, the ensemble states separated into two clusters, denoting increased similarity between spiking patterns associated with each object (Fig. 5, phase 2). Following the presentation of the grip cue, SSIMS representations shifted, forming a new set of clusters reflecting grip type in addition to object similarity. In two of the monkeys (G and S), this resulted in a set of four independent clusters, each associated with one grip/object pairing that persisted beyond the instructed delay period and into the action execution (contact) phase. Note that the two power grips formed distinct but neighboring clus-

ters. In the third monkey, grip-based clustering overshadowed object-related influences, generating only three clusters in the later stage of the task. Notably, for Monkey R the power grip formed a single cluster regardless of which object was the target. This observation is in agreement with the apparent lack of object-selective neurons observed in the later parts of the task for Monkey R (Fig. 4).

**Decoding object type and grip strategy**

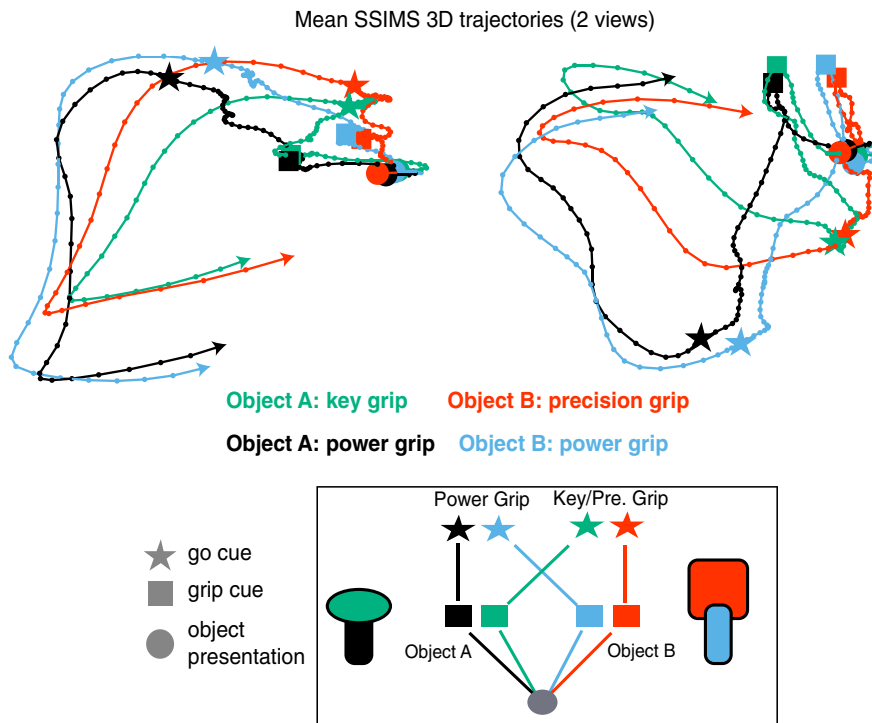
We next used decoding methods to quantify the information related to objects and grips present in PMv ensemble activity. These results also provide a direct measure of useful information that could be extracted from PMv for potential neuroprosthetic control purposes. We used SSIMS projections for partially overlapping 1 s time windows (advancing in increments of 50 ms) to provide a continuous estimate of the ensemble states throughout the task. For classification purposes, 15-dimensional SSIMS projections were used. Decoding was performed using nearest neighbor classifiers tested using leave-one-out cross validation. Classification results across trial time for object and grip type are shown in Figure 6. For all monkeys, mean performance for a given type of classification was below the corresponding 99% confidence limit of the chance distribution before the trial started (evaluated using a Monte Carlo approach with 10,000 random shuffles of the trial labels); this could be considered as the baseline state. In agreement with our initial clustering analysis, object classification rose above the 99% confidence interval after the target object was illuminated, well in advance of significant grip-related



**Figure 6.** Decoding object and grip type. A nearest-neighbor classifier using leave-one-out cross validation was used to classify both grip type and object strategy at each time point based on SSIMS projections generated using partially overlapping 1 s windows (50 ms shifts). Results for each monkey are shown separately (mean over 4 sessions in each animal, with shading denoting SD). The upper bounds for the 99% confidence intervals (dashed lines) for object/grip classification were determined using 10,000 random shuffles of the trial labels (values differ because there are 3 grips and 2 objects).

decoding. That is, the object type being viewed could be categorized based on ensemble firing differences on single trials. In Monkeys G and S, the accuracy of object classification continued to increase up to the moment of action execution, reaching >90% around the time of object contact. This finding indicates that the network continued to carry information sufficient to clearly distinguish the two objects. A different pattern was observed in Monkey R, where object classification

dropped to near-chance levels after an early peak following object presentation (Fig. 6). This failure to accurately classify objects across the instructed delay is consistent with SSIMS projections showing a single cluster common to both power grips in Monkey R (Fig. 5). In all monkeys, grip classification accuracy gradually increased during the delay period following the grip instruction cue, with near perfect (>95% correct) classification by the time of contact with the object.



**Figure 7.** SSIMS trajectories: linking objects to actions. The temporal evolution of ensemble activity is tracked by generating SSIMS projections from sliding time windows (1 s duration, sliding by 50 ms). The resulting space captures the relationship between activity patterns observed at different task phases. The mean trajectories for the four different types of trials are shown in different colors for one session in Monkey G (two viewpoints are provided to display 3D structure). Dots along the lines mark the 50 ms shifts in the time bins. Trajectories were smoothed with a Gaussian kernel (400 ms SD). Object presentation events, grip cues, and go cues, are highlighted by circles, stars, and squares, respectively. The graph in the boxed region shows a summary of the general pattern observed. Similar results were obtained for the other two monkeys.

**Temporal evolution of neural ensemble states**

We generated three-dimensional SSIMS trajectories to evaluate the temporal evolution of ensemble activity patterns across task phases (see Materials and Methods for details). Whereas the SSIMS projections from a single-task phase (Fig. 5) display the relationship between different ensemble states at a given time, the SSIMS trajectories highlight the relationship between the activity patterns observed at different time points in the task. The decoding analysis presented in Figure 6 quantifies the separation of activity patterns related to different object + grip combinations at different time points; the mean neural trajectories in Figure 7 provide an intuitive way to visualize the convergence/divergence of the ensemble activity patterns through the course of object presentation, action selection, and movement execution. This illustration highlights how the ensemble states associated with a given object are initially most similar, resulting in two object-related clusters. Following the grip instruction, however, neural activity gradually shifts into a new pattern where ensemble states associated with specific grip strategies are most similar, consistent with their engagement in computing grip.

**Discussion**

We evaluated the evolution of ensemble activity in primate PMv as visual information from objects and grip instruction cues was used to plan and execute grasping movements. Monkeys would first view an object affording two possible grasping strategies and then receive a light cue instructing how it should be grasped. It is important to note that the monkeys had to rely on two different types of instructions to determine the correct grip to use. Object illumination allowed the monkeys to view the 3D shape of the target object. This has been shown to involve parietal areas (in

particular AIP; Luppino et al., 1999; Murata et al., 2000). Next, the monkeys had to interpret the color of the grip cue light to choose the correct grip. This type or arbitrary stimulus–response association is known to preferentially engage frontal regions, such as the supplementary motor area and dorsal premotor cortex. Thus, the task required the combination of two qualitatively different types of visually derived information engaging largely segregated cortical circuits: (1) information directly related to object geometry, and (2) information related to the grip instruction through a learned stimulus–response association. Clustering of neural activity patterns at different task phases demonstrates a transition from a set of ensemble states most closely related to objects to a different set reflecting object + grip combinations. Our results show that ensemble activity recorded within a 4 × 4 mm area of PMv contains information related to the full transformation leading from object detection to action execution, dynamically responding to information as it becomes available.

The model of PMv functional organization proposed by Fagg and Arbib (1998), based on the seminal work by Rizzolatti et al. (1988), has remained relatively unchanged in recent years (Rizzolatti and Luppino, 2001; Raos et al., 2006; Carpaneto et al., 2011). This model

proposes that subcircuits in PMv represent archetypal grip strategies comprising an elementary “movement vocabulary” activated by visually driven activity in parietal cortex. When multiple grips are possible, the model suggests a “winner take all” selection process that ultimately leads to the activation of a circuit representing a single canonical grip. Previous experiments have shown accurate decoding of many different grip types, but have not explicitly examined interactions between grip affordances for individual objects: they either used a task in which objects were associated with a single type of grip, or presented objects and grip instructions simultaneously (Fluet et al., 2010; Townsend et al., 2011; Schaffelhofer et al., 2015). By contrast, in the present experiment we compared both single neurons and ensemble responses to objects that can be grasped in different ways, separating in time the visual presentation of the target object from the instruction of how to grasp it. As a particular grasping strategy was selected from a set of possible movements associated with a given object, we observed a shift from a condition where ensemble states related to the same object were most similar, to a regime where the planned/executed grip was the main factor determining the relationship among ensemble states. Notably, we observed clusters of ensemble states associated with each object before the grip was specified. These neural activity patterns diverged for the two objects well in advance of the grip cue, which was unknown to the monkey when the object was presented. Activity during the object presentation stage did not appear to represent “default” grip strategies, because the ensemble states in question were different from any of those engaged after the grip instruction was provided. These “object-related” ensemble states



appear to be stepping-stones that bring the network closer to the desired output state (i.e., one that generates downstream activity triggering the appropriate grasp) as part of an ongoing dynamic process. Preparatory activity (preceding movement) can be mathematically described as the early phases in the evolution of a complex dynamical system that will eventually reach the desired output state (Churchland et al., 2010). Under this view, “representing” objects and/or grips as such is not necessary (or even desirable) as long as the appropriate descending cortical outputs are generated. Our results are fully compatible with this formulation: the ensemble appears to continuously incorporate information throughout the instructed delay, gradually shifting the population response toward the correct output patterns.

Our results show that presenting objects with different combinations of grip affordances elicited distinct activity patterns in PMv. Furthermore, in two of our three NHPs (G and S) we observed statistically separable patterns of activity associated with engaging power grips aimed at different objects (even though the relevant parts of the objects were of exactly the same dimensions and positions relative to the monkeys). These distinct activity patterns emerged as soon as the correct grip type was cued (before movement was initiated), and persisted throughout the execution of the grasping movement. Schaffelhofer et al. (2015) recently demonstrated that specific grips can be reliably classified even when aimed at different objects. It is important to note that, for that study, monkeys executed grasping movements in the dark. It is likely that visual input plays a key role in the object-related effects we observe. Additionally, in the present study the SSIMS projections for power grip trials aimed at different objects formed separate but adjacent clusters for Monkeys G and S, showing that the associated neural activity patterns, although distinct, still tended to be similar (Fig. 5). A decoder therefore should still be able to generalize across the two types of power grips, reproducing the decoding results presented by Schaffelhofer et al. (2015).

We observed a different pattern in Monkey R: in this case, the power grip clusters for the two objects overlapped, resulting in relatively poor object decoding. This may reflect a difference in the neuronal subpopulation sampled. It is possible that, in this monkey, a larger portion of PMv was buried in the bank of the arcuate sulcus, beyond the reach of the microelectrode arrays used in this study. In this case, recording electrodes may have encroached upon more caudal MI populations. The decrease in decoding accuracy for both object and grip information during the instructed delay periods is consistent with this scenario. Alternatively, Monkey R may have approached the task using a different cognitive strategy from the other two NHPs. Using a single activity pattern to generate power grips for both objects could be viewed as an example of successful behavioral generalization. Generalization across different contexts is a characteristic of highly encephalized species that has been notoriously hard to replicate in artificial intelligence systems (Kelley and Long, 2010). Monkeys can be highly variable in their ability to generalize (Hikosaka et al., 1995). Merging different neural states with equivalent motor outputs into a single activity pattern may reflect a mechanism of generalization used by the motor system when learning new tasks. Further experiments with independent measures of generalization will be required to explore this possibility.

Our results demonstrate that very similar grasping movements can be related to different (that is, statistically separable) neural activity patterns when engaged within the “context” of different objects. Given the high-dimensional nature of neural ensemble states, there are likely to be multiple states representing correct solutions to a given grasping problem: that is, there are

likely to be many different cortical network activity patterns that will ultimately lead to the same downstream output to  $\alpha$  motor neurons. A many-to-one mapping with multiple “output-equivalent” network states that generate the same behavioral response could simplify the process of mapping sensory inputs to desired motor outputs, in effect expanding the set of target states that provide a correct solution to a given behavioral challenge. This organizational principle could also facilitate the ability to trigger a particular behavior in response to different types of stimuli, since it would circumvent the problem of funneling different kinds of information into a single state that uniquely triggers the desired motor output. The experience-driven tuning of this system would result in “affordance landscapes” unique to each subject, dynamically linking visual inputs with motor responses using multiple possible pathways through the high-dimensional space of possible network states. Note that these output-equivalent states could potentially hold different types of information not reflected in the motor output: for example, two states ultimately leading to power grips could also reflect other potential grip affordances present in a given context. Preserving information about other possible grips in addition to the one being currently engaged could provide some ethological advantage (if a rapid change in grasping strategy is required, for example). Examining grasping-related neural activity over a broader range of behavioral tasks using more detailed kinematic measurements will be necessary to determine how PMv represents similar movements performed in different contexts, and to what degree subtle differences in movement execution are reflected in neural activity.

Neurons in ventral premotor cortex have also been shown to respond to somatosensory stimuli. Remarkably, many neurons display dual visual/tactile receptive fields organized in a body-anchored reference frame; this kind of neurons respond to visual stimuli approaching a specific body part (Rizzolatti et al., 1981; Graziano et al., 1994; Fogassi et al., 1996). We observed the highest grip and object decoding accuracy around the time monkeys made contact with the objects (Fig. 5). It is likely that the activity patterns we describe incorporate additional visual and somatosensory components as the objects enter the monkey’s peripersonal space. Further experiments will be needed to isolate the contribution of this type of visual/somatosensory signals to activity in PMv.

### Brain–computer interface applications

Brain–computer interfaces (BCIs) allow people with compromised mobility to use neural activity as a command signal for assistive devices (Hochberg et al., 2006; 2012; Collinger et al., 2013). Our results in monkeys suggest that premotor areas may provide useful signals for BCI applications. We have demonstrated robust decoding of both objects and grips, and have shown that object-related activity can be present well in advance of the emergence of an action plan. Decoding these activity patterns could provide advanced information about possible action targets to assistive devices. Our data also suggests that some degree of contextual information is present in PMv, allowing us to distinguish between power grips aimed at different objects. Wodlinger et al. (2015) have recently reported that BCI kinematic decoders calibrated in the absence of objects may fail to perform well once objects are introduced. We conclude that “seeing an object elicits a component or factor” which interferes with direct decoding of kinematics, disrupting intended grasping commands and causing the robotic limb under neural control to back away from the object, contrary to the stated intention of the

user. Remarkably, the effect disappeared once the participant closed her eyes. This finding is in agreement with our observations, and suggests that object-related information is present in human motor cortex. Taking into account this type of context-related signals may be a critical step in achieving naturalistic, generalizable, BCI control.

## References

- Barrese JC, Rao N, Paroo K, Triebwasser C, Vargas-Irwin C, Franquemont L, Donoghue JP (2013) Failure mode analysis of silicon-based intracortical microelectrode arrays in non-human primates. *J Neural Eng* 10:066014. [CrossRef Medline](#)
- Belmalih A, Borra E, Contini M, Gerbella M, Rozzi S, Luppino G (2009) Multimodal architectonic subdivision of the rostral part (area F5) of the macaque ventral premotor cortex. *J Comp Neurol* 512:183–217. [CrossRef Medline](#)
- Brochier T, Umiltà MA (2007) Cortical control of grasp in non-human primates. *Curr Opin Neurobiol* 17:637–643. [CrossRef Medline](#)
- Carpaneto J, Umiltà MA, Fogassi L, Murata A, Gallese V, Mìcera S, Raos V (2011) Decoding the activity of grasping neurons recorded from the ventral premotor area F5 of the macaque monkey. *Neuroscience* 188:80–94. [CrossRef Medline](#)
- Churchland MM, Yu BM, Sahani M, Shenoy KV (2007) Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Curr Opin Neurobiol* 17:609–618. [CrossRef Medline](#)
- Churchland MM, Cunningham JP, Kaufman MT, Ryu SI, Shenoy KV (2010) Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron* 68:387–400. [CrossRef Medline](#)
- Collinger JL, Wodlinger B, Downey JE, Wang W, Tyler-Kabara EC, Weber DJ, McMorland AJ, Velliste M, Boninger ML, Schwartz AB (2013) High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet* 381:557–564. [CrossRef Medline](#)
- Fagg AH, Arbib MA (1998) Modeling parietal-premotor interactions in primate control of grasping. *Neural Netw* 11:1277–1303. [CrossRef Medline](#)
- Fluet MC, Baumann MA, Scherberger H (2010) Context-specific grasp movement representation in macaque ventral premotor cortex. *J Neurosci* 30:15175–15184. [CrossRef Medline](#)
- Fogassi L, Gallese V, Fadiga L, Luppino G, Matelli M, Rizzolatti G (1996) Coding of peripersonal space in inferior premotor cortex (area F4). *J Neurophysiol* 76:141–157. [Medline](#)
- Fogassi L, Gallese V, Buccino G, Craighero L, Fadiga L, Rizzolatti G (2001) Cortical mechanism for the visual guidance of hand grasping movements in the monkey: a reversible inactivation study. *Brain* 124:571–586. [CrossRef Medline](#)
- Godschalk M, Lemon RN, Nijs HG, Kuypers HG (1981) Behaviour of neurons in monkey peri-arcuate and precentral cortex before and during visually guided arm and hand movements. *Exp Brain Res* 44:113–116. [Medline](#)
- Graziano MS, Yap GS, Gross CG (1994) Coding of visual space by premotor neurons. *Science* 266:1054–1057. [CrossRef Medline](#)
- Hikosaka O, Rand MK, Miyachi S, Miyashita K (1995) Learning of sequential movements in the monkey: process of learning and retention of memory. *J Neurophysiol* 74:1652–1661. [Medline](#)
- Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442:164–171. [CrossRef Medline](#)
- Hochberg LR, Bacher D, Jarosiewicz B, Masse NY, Simeral JD, Vogel J, Haddadin S, Liu J, Cash SS, van der Smagt P, Donoghue JP (2012) Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485:372–375. [CrossRef Medline](#)
- Kelley T, Long L (2010) Deep blue cannot play checkers: the need for generalized intelligence for mobile robots. *J Robotics* 2010:523757. [CrossRef](#)
- Kurata K, Tanji J (1986) Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements. *J Neurosci* 6:403–411. [Medline](#)
- Lehky SR, Sereno ME, Sereno AB (2013) Population coding and the labeling problem: extrinsic versus intrinsic representations. *Neural Comput* 25:2235–2264. [CrossRef Medline](#)
- Lemon RN (1993) The G. L. Brown Prize Lecture: cortical control of the primate hand. *Exp Physiol* 78:263–301. [CrossRef Medline](#)
- Luppino G, Murata A, Govoni P, Matelli M (1999) Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp Brain Res* 128:181–187. [CrossRef Medline](#)
- Murata A, Fadiga L, Fogassi L, Gallese V, Raos V, Rizzolatti G (1997) Object representation in the ventral premotor cortex (area F5) of the monkey. *J Neurophysiol* 78:2226–2230. [Medline](#)
- Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol* 83:2580–2601. [Medline](#)
- Napier JR (1980) *Hands*. New York: Pantheon Books.
- Nelissen K, Vanduffel W (2011) Grasping-related functional magnetic resonance imaging brain responses in the macaque monkey. *J Neurosci* 31:8220–8229. [CrossRef Medline](#)
- Raos V, Umiltà MA, Murata A, Fogassi L, Gallese V (2006) Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey. *J Neurophysiol* 95:709–729. [CrossRef Medline](#)
- Rizzolatti G, Luppino G (2001) The cortical motor system. *Neuron* 31:889–901. [CrossRef Medline](#)
- Rizzolatti G, Scandolara C, Matelli M, Gentilucci M (1981) Afferent properties of periarculate neurons in macaque monkeys: I. Somatosensory responses. *Behav Brain Res* 2:125–146. [CrossRef Medline](#)
- Rizzolatti G, Camarda R, Fogassi L, Gentilucci M, Luppino G, Matelli M (1988) Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Exp Brain Res* 71:491–507. [CrossRef Medline](#)
- Schaffelhofer S, Agudelo-Toro A, Scherberger H (2015) Decoding a wide range of hand configurations from macaque motor, premotor, and parietal cortices. *J Neurosci* 35:1068–1081. [CrossRef Medline](#)
- Suner S, Fellows MR, Vargas-Irwin C, Nakata GK, Donoghue JP (2005) Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. *IEEE Trans Neural Syst Rehabil Eng* 13:524–541. [CrossRef Medline](#)
- Tanné-Gariépy J, Rouiller EM, Boussaoud D (2002) Parietal inputs to dorsal versus ventral premotor areas in the macaque monkey: evidence for largely segregated visuomotor pathways. *Exp Brain Res* 145:91–103. [CrossRef Medline](#)
- Theys T, Pani P, van Loon J, Goffin J, Janssen P (2012) Selectivity for three-dimensional shape and grasping-related activity in the macaque ventral premotor cortex. *J Neurosci* 32:12038–12050. [CrossRef Medline](#)
- Theys T, Pani P, van Loon J, Goffin J, Janssen P (2013) Three-dimensional shape coding in grasping circuits: a comparison between the anterior intraparietal area and ventral premotor area F5a. *J Cogn Neurosci* 25:352–364. [CrossRef Medline](#)
- Townsend BR, Subasi E, Scherberger H (2011) Grasp movement decoding from premotor and parietal cortex. *J Neurosci* 31:14386–14398. [CrossRef Medline](#)
- Umiltà MA, Brochier T, Spinks RL, Lemon RN (2007) Simultaneous recording of macaque premotor and primary motor cortex neuronal populations reveals different functional contributions to visuomotor grasp. *J Neurophysiol* 98:488–501. [CrossRef Medline](#)
- van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9:2579–2605.
- Vargas-Irwin C, Donoghue JP (2007) Automated spike sorting using density grid contour clustering and subtractive waveform decomposition. *J Neurosci Methods* 164:1–18. [CrossRef Medline](#)
- Vargas-Irwin CE, Brandman DM, Zimmermann JB, Donoghue JP, Black MJ (2015) Spike train SIMilarity space (SSIMS): a framework for single neuron and ensemble data analysis. *Neural Comput* 27:1–31. [CrossRef Medline](#)
- Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J Neurophysiol* 76:1310–1326. [Medline](#)
- Wodlinger B, Downey JE, Tyler-Kabara EC, Schwartz AB, Boninger ML, Collinger JL (2015) Ten-dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations. *J Neural Eng* 12:016011. [CrossRef Medline](#)