Reach and gaze representations in macaque parietal and premotor grasp areas

Lehmann, S J; Scherberger, H

Abstract: Voluntary movements are frequently composed of several actions that are combined to achieve a specific behavior. For example, prehension involves reaching and grasping actions to transport the hand to a target to grasp or manipulate it. For controlling these actions, separate parietofrontal networks have been described for generating reaching and grasping actions. However, this separation has been challenged recently for the dorsomedial part of this network (area V6A). Here we report that the anterior intraparietal (AIP) and the rostral ventral premotor area (F5) in the macaque, which are both part of the dorsolateral parietofrontal network and causally linked to hand grasping movements, also represent spatial information during the execution of a reach-to-grasp task. In addition to grip type information, gaze and target positions were represented in AIP and F5 and could be readily decoded from single unit activity in these areas. Whereas the fraction of grip type tuned units increased toward movement execution, the number of cells with spatial representations stayed relatively constant throughout the task, although more prominently in AIP than in F5. Furthermore, the recorded target position signals were substantially encoded in retinotopic coordinates. In conclusion, the simultaneous presence of grasp-related and spatial information in AIP and F5 suggests at least a supportive role of these spatial signals for the planning of grasp actions. Whether these spatial signals in AIP and F5 also play a causal role for the planning of reach actions would need to be the subject of further investigations.

DOI: https://doi.org/10.1523/JNEUROSCI.5568-12.2013

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-91146
Published Version

Originally published at:
DOI: https://doi.org/10.1523/JNEUROSCI.5568-12.2013
Voluntary movements are frequently composed of several actions that are combined to achieve a specific behavior. For example, pre-hension involves reaching and grasping actions to transport the hand to a target to grasp or manipulate it. For controlling these actions, separate parietofrontal networks have been described for generating reaching and grasping actions. However, this separation has been challenged recently for the dorsomedial part of this network (area V6A). Here we report that the anterior intraparietal (AIP) and the rostral ventral premotor area (F5) in the macaque, which are both part of the dorsolateral parietofrontal network and causally linked to hand grasping movements, also represent spatial information during the execution of a reach-to-grasp task. In addition to grip type information, gaze and target positions were represented in AIP and F5 and could be readily decoded from single unit activity in these areas. Whereas the fraction of grip type tuned units increased toward movement execution, the number of cells with spatial representations stayed relatively constant throughout the task, although more prominently in AIP than in F5. Furthermore, the recorded target position signals were substantially encoded in retinotopic coordinates. In conclusion, the simultaneous presence of grasp-related and spatial information in AIP and F5 suggests at least a supportive role of these spatial signals for the planning of grasp actions. Whether these spatial signals in AIP and F5 also play a causal role for the planning of reach actions would need to be the subject of further investigations.

Introduction

Humans and monkeys perform reaching and grasping movements that are spatially and temporally well coordinated (Jeannerod et al., 1995). Planning and execution of such limb movements occur in separate frontoparietal networks for reaching (dorsomedial stream) and grasping (dorsolateral stream) (Rizzolatti and Luppino, 2001; Culham et al., 2006). Object manipulations are furthermore strongly linked to eye movements (Johansson et al., 2001; Neggers and Bekkering, 2002), which are controlled by yet another frontoparietal network including the frontal eye field and lateral intraparietal cortex (Ferraina et al., 2002; Schiller and Tehovnik, 2005). These distinct frontoparietal networks smoothly accomplish dexterous object manipulations by coordinating looking, reaching, and grasping actions.

The posterior parietal cortex is known to process visual information for the guidance of actions (Hyvarinen and Poranen, 1974; Mountcastle et al., 1975; Jeannerod et al., 1995). Specifically, the anterior intraparietal area (AIP) is relevant for shaping the hand to grasp (Taira et al., 1990; Sakata et al., 1995; Sakata et al., 1997; Murata et al., 2000). AIP represents visual object properties (Rizzolatti and Luppino, 2001) and context-specific hand-grasping movements (Baumann et al., 2009). It has major connections with many parietal areas, indicating its important role for sensory integration, and is strongly and reciprocally connected with the rostral ventral premotor area F5 (dorsolateral stream for grasping) (Matelli et al., 1986; Luppino et al., 1999; Tammé-Gariépy et al., 2002; Borra et al., 2008).

The ventral premotor cortex (PMv), in particular its latero-rostral portion (area F5), similarly participates in visuomotor transformations for grasping (Gentilucci et al., 1983; Murata et al., 1997; Cerri et al., 2003; Umiltà et al., 2007). F5 projects to the hand area of primary motor cortex (Matsumura and Kubota, 1979; Muakkassa and Strick, 1979; Matelli et al., 1986; Dum and Strick, 2005) and to the spinal cord (Dum and Strick, 1991; He et al., 1993; Borra et al., 2010), implying a direct involvement in cortical motor output for hand grasping. Functionally, F5 neurons represent specific grip types both during motor planning and execution (Murata et al., 1997; Kakei et al., 2001; Raos et al., 2006; Stark et al., 2007; Fluet et al., 2010).

These areas are causally relevant for their respective functions (Rizzolatti and Luppino, 2001; Buneo and Andersen, 2006). Inactivation of AIP or F5 leads to a specific grasping deficit, whereas reaching and eye movements are rather unimpaired (Gallese et al., 1994; Fogassi et al., 2001). However, reach target and gaze position signals have been described in F5, but not in AIP (Boussaoud et al., 1993; Mushiake et al., 1997; Stark et al., 2007). Furthermore, a strict separation of reach and grasp representa-
Reach and gaze position signals in AIP and less frequently in F5, we found strong action networks by investigating the representation of spatial positions has been challenged for the dorsomedial area V6A (Fattori et al., 2005; Fattori et al., 2009).

Materials and Methods

**Experimental setup.** Two purpose-bred female rhesus monkeys (Macaca mulatta) participated in this study (animals P and S; weight 4.5 and 5.5 kg, respectively). They were pair-housed in a spacious and enriched environment. All procedures and animal care were conducted in accordance with the regulations set by the Veterinary Office of the Canton of Zurich, the guidelines for the care and use of mammals in neuroscience and behavioral research (National Research Council, 2003), and in agreement with German and European laws governing animal care.

Animals were habituated to comfortably sit upright in an individually adjusted primate chair with the head rigidly fixed to the chair. A grasp target was located at a distance of 24 cm in front of the animal. The target consisted of a handle that could be grasped with two different grip types, either with a precision grip (using index finger and thumb in opposition) or a whole-hand power grip (Baumann et al., 2009; Fluet et al., 2010). In addition to the straight-ahead position, the location of the target could be moved by two motors horizontally (on a circular pathway) and vertically, thus allowing five different target positions with spacing of 11 cm (24.5° visual angle). Left and right off-center panels were tilted in correspondence with the horizontal pathway (Fig. 1C).

Similarly, eye fixation position was varied by addressing different LEDs, either in combination with target position variation or independent of it. Spatial target position was changed automatically during the intertrial period in the dark; however, the sound of the moving motors could potentially provide some clues to the animal about the upcoming target location before the start of the next trial. To illuminate the handle in the dark, two dedicated spotlights were positioned to the left and right of the handle (outside of the animal’s reach). Eye position was monitored with an optical eye tracking system (ET-49B; Thomas Recording). Touch sensors in front of the animal’s hips were used to monitor the hand-resting position for both hands. Animal behavior and all stimulus parameters were controlled in LabView Realtime (National Instruments) with a time resolution of 5 ms using custom-written software.

**Task paradigms.** Monkeys were trained to perform a delayed reach-to-grasp task, in which they were instructed to grasp a target with either a power grip or a precision grip. Animals initialized each trial by placing both hands on the hand rest sensors and fixing a red LED while otherwise sitting in the dark. The trial started with a baseline (fixation) epoch (500–700 ms), during which the animal had to maintain its resting position in the dark. In the following cue epoch (fixed length: 800 or 1000 ms), an additional LED was shown close to the fixation position that informed the animal about the required grasp type (green LED: power grip, orange LED: precision grip). At the same time,
the grasp target was illuminated, which revealed the handle position in space. In the following planning epoch of variable length (800–1200 ms), only the fixation light was visible and the animal could plan, but not yet execute the movement. A short blink of the fixation light (the "go" cue) then instructed the animal to reach and grasp the target (movement epoch) with its left arm (contralateral to the recording chamber). Planning and movement epochs were in complete darkness except for the red LED light that the animal had to keep fixating on throughout the task (window radius: 11.4°). In this paradigm, we therefore separated in time the representation of visual stimuli from motor planning and execution. Based on previous studies, we also expect that motor activity is similarly represented in AIP and F5 when grasping in the light or dark (Murata et al., 2000; Raos et al., 2006).

All correctly executed trials were rewarded with a fixed amount of juice, and the animal could initiate the next trial after a short intertrial interval. Error trials were immediately aborted without reward. To maintain a high motivation for obtaining fluid rewards, animals were restricted from access to water before training and recording sessions.

In different task conditions, we systematically varied target position and gaze position during the intertrial interval, which were grouped in three different subtasks:

1. Combined gaze-and-reach variation (CV): target and gaze position were varied together, both at the center or at the left, right, top, or bottom position (Fig. 1D, left).
2. Target variation (TV): gaze position was located at the center of the workspace, while target position was either at the center or at the left, right, top, or bottom position (Fig. 1D, middle).
3. Gaze variation (GV): target position was located at the center of the workspace while gaze position was varied between the center, left, right, top, and bottom position (Fig. 1D, right).

In combination with the two different grip types, this led to a set of 26 task conditions that were presented pseudo-randomly interleaved in blocks of typically 10 trials per condition.

**Surgical procedures and MRI scans.** Details of the surgical procedures and MRI scans have been described previously (Baumann et al., 2009; Fluet et al., 2010). In short, a titanium head post was secured in a dental acrylic head cap and a custom made oval-shaped recording chamber (material PEEK (polyether ether ketone); outer dimensions, 40 mm × 30 mm). A bank of parietal sulcus (Borra et al., 2008), whereas in F5 we recorded structures (Fig. 1). AIP was then defined as the rostral part of the lateral principal sulcus (Belmalih et al., 2009).

**Neuronal recordings.** Single-unit (spiking) activity was recorded using quartz-glass-coated platinum/tungsten electrodes (impedance 1–2 MΩ at 1 kHz) that were positioned simultaneously in AIP and F5 by two five-channel micromanipulators (Mini-Matrix, Thomas Recording). Neural signals were amplified (400×), digitized with 16-bit resolution at 30kS/s using a Cerebus Neural Signal processor (Blackrock Microsystems), and stored on a hard drive together with the behavioral data.

**Data analysis.** All data analysis was performed offline. Neural signals were high-pass filtered with a cutoff frequency of 500 Hz and single units were isolated using principal component analysis techniques (Offline Sorter v2.8.8, Plexon). Using Matlab (Mathworks) for further analysis, we included all units in our database that had an average firing rate of at least 5 Hz in one of the task conditions and that were stably recorded in at least 7 trials (typically 10 trials) per condition (182–260 trials in total).

Peristimulus time histograms were generated using a gamma distribution as a causal kernel (parameters: shape α = 1.5, rate β = 30; Baumann et al., 2009). However, all statistical tests were based on exact spike counts.

The preferred and nonpreferred grip type and the preferred and nonpreferred spatial position in the CV, TV, and GV tasks were determined for each cell from the mean activity in the time interval from fixation onset to movement end. Activity was averaged across all trials of the same grip type or spatial position. The grip type with the higher (or lower) mean firing rate was defined as the preferred (or nonpreferred) grip type. Similarly, the spatial position with the highest (or lowest) mean firing rate for a given subtask was defined as the preferred (or nonpreferred) position.

To test the significance of tuning for grip type and for spatial position in each task epoch (fixation, cue, planning, and movement), we calculated the mean firing rate in every trial (spike count/length of epoch) and performed a two-way ANOVA (factors grip type and spatial position; p < 0.01) separately for each epoch.

Tuning onset of each cell was determined with a sliding window ANOVA analysis for grip type and spatial position. Here, a two-way ANOVA (with a fixed p-value of 0.01) was repeated for a series of windows of 200 ms length that were shifted in time steps of 50 ms. Tuning onset was then defined as the first occurrence of (at least) five consecutive windows with a significant ANOVA. This definition prevented a possible bias due to multiple testing. The analysis was performed for three different alignments of spiking activity: to fixation onset, cue offset, and onset of movement epoch (the "go" cue).

A similar approach was used to check the quality of tuning for each neuron by analyzing the discriminability of different conditions, as quantified by the receiver operator characteristic (ROC) score (Townsend et al., 2011). For this, we calculated a ROC score from the trials of the two grip types and from the preferred and nonpreferred position in each task: CV, TV, and GV. For each grip type or spatially tuned unit, the ROC score was calculated in a sliding window of 200 ms width, shifted in steps of 20 ms. The population average of the resulting ROC curves over time were then compared for grip type and the various spatial conditions in the CV, TV, and GV task (one-way ANOVA, p < 0.01).

To further investigate the tuning of individual neurons, we modeled the firing rates of each neuron (in specific task epochs) in a stepwise linear model including the factors grip type (GT), target position (T), and gaze position (G):

\[ f = a + gt \cdot GT + t \cdot T + g \cdot G \]

(1)

Because the spatial factors T and G each have a horizontal (x) and a vertical (y) component, a more detailed description of the model would be:

\[ f = a + gt \cdot GT + t_x \cdot T_x + t_y \cdot T_y + g_x \cdot G_x + g_y \cdot G_y \]

(2)

By starting with the constant model: \( f = a \), additional components were added in a stepwise fashion until no further significant improvements could be obtained (MATLAB function: stepwisefit; \( p < 0.05 \)). This allowed us to categorize each neuron according to its significant modulations. A neuron was considered modulated by a spatial factor if the model contained either a significant horizontal or vertical component. This resulted in a spatial categorization of neurons that were modulated by target position, gaze position, or by neither or both factors.

In addition, the coefficients of the linear regression analysis were further processed by calculating the angular difference between the vectors \( t = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \) and \( g = \begin{pmatrix} g_x \\ g_y \end{pmatrix} \) as well as the length contrast (LC) between both vectors:

\[ LC = \frac{||t|| - ||g||}{||t|| + ||g||} \]

(3)

where \( ||t|| \) and \( ||g|| \) describe the length of vectors \( t \) and \( g \), respectively.

Neurons were considered to encode spatial information retinotopically, if: (1) they were both significantly modulated by target and gaze position (as revealed by the stepwise fit), (2) they showed similar vector lengths for \( t \) and \( g \) (\( -0.33 < LC < +0.33 \); i.e., vector lengths differed by less than factor 2), and (3) the angular difference between \( t \) and \( g \) was at least 135°, corresponding to \( t \) and \( g \) pointing approximately in opposite directions. These assumptions are reasonable, because for retinotopic coding \( t = -g \) and thus \( f = a + gt \cdot GT + t \cdot (T - G) \).
For neural decoding, we used a maximum likelihood estimation approach (Scherberger et al., 2005; Townsend et al., 2011). We simulated the decoding of grip type, individual spatial conditions, and the spatial target position, retinotopic target position, and gaze position from neural activity during the different task epochs. The decoding simulation was performed based on the sequentially recorded populations of neurons in AIP and F5, which for this analysis were assumed to be recorded simultaneously (hence the term: decoding simulation). Decoding was performed with all units that were significantly modulated for the conditions to be analyzed in the respective task epoch (one-way ANOVA, \( p < 0.05 \)). We calculated the percentage of correctly decoded conditions from 100 simulated repetitions and iterated this process 200 times to obtain a mean performance and an SD, from which 95% confidence limits were inferred by assuming a normal distribution. To illustrate the decoding results, we plotted the frequency of instructed and decoded condition pairs in a color-coded confusion matrix, in which correctly classified trials line up on the diagonal.

**Results**

Two female macaque monkeys (animals P and S) performed a delayed reach-to-grasp task, in which a target placed at different spatial locations had to be grasped either with a precision grip or a power grip while eye position was controlled (see Materials and Methods). Target location and grip type instruction were disclosed during a cue epoch, but animals had to withhold movement execution until a go signal was presented (Fig. 1A). Spatial target position and gaze position were systematically varied (Fig. 1C,D). Target and gaze position were varied together (CV task; Fig. 1D, left), only the target position was varied while eye position was held constant at the center (TV task; Fig. 1D, middle), or the grasp target remained at the center and gaze was varied (GV task; Fig. 1D, right). In combination with two grip types, this led to a total of 26 task conditions.

**Task behavior**

Animals made approximately 600–1000 correct trials per session and performed the task with high accuracy (e.g., error rates due grip type confusion < 3%). Task behavior was consistent across the different task conditions with a mean reaction time of 270 ms (SD: 40 ms) for animal P and 250 ms (SD: 40 ms) for animal S, which was also independent of grip type. Movement times were slower for precision grips (animal P: 350 ms, SD: 160 ms; animal S: 450 ms, SD: 120 ms) than power grips (animal P: 300 ms, SD: 70 ms; animal S: 350 ms, SD: 70 ms) and were slightly faster for targets closer to the hand starting position (left and down targets) than further away. However, there was no systematic difference between central and peripheral, or foveated and nonfoveated targets. In addition, even though we did not measure hand kinematics explicitly, we observed no differences of finger aperture or forearm rotation as a function of target location.

**Tuning for grip type, reach, and gaze position**

We recorded a total of 353 single units in AIP (animal P: 207 units, animal S: 146 units) and 585 units in F5 (P: 284 units; S: 301 units). A majority of the recorded cells were modulated by the reach-to-grasp task, either exclusively for grip type, target, or gaze position or in combination. Recording tracks are charted in Figure 1E. By projecting the recording positions of the AIP and F5 cells on an axis that was approximately parallel to the intraparietal sulcus and the relevant part of the arcuate sulcus, respectively, we found that the cortical distributions of grip-type and spatially tuned neurons in AIP and F5 were approximately uniform, suggesting that both groups of neurons were anatomically intermingled. Figure 2 shows an example neuron from F5 that was strongly tuned for the instructed grip type and for target and gaze position. For precision grips, spiking activity was clearly higher during the planning epoch and toward movement execution (Fig. 2A–C) than for power grips (Fig. 2D–F). In addition, the activity of this neuron was strongly modulated by the spatial target and gaze position, in particular during precision grip trials (Fig. 2A–C). In the CV task, activity differences were only moderate, with the strongest modulation between the top and bottom positions (Fig. 2A). Position modulation was stronger in the TV and GV task, when target and gaze positions were spatially separated (Fig. 2B, C). Here, modulation by target and gaze position was not independent, but rather consistent with a retinotopic representation of the grasp target. In fact, activity during the cue and planning epoch was highest for target positions to the left and down of the eye fixation (gaze) position independent of whether the target position was varied and gaze kept constant (TV task; Fig. 2B, cyan and red curves) or gaze position was varied and the target position kept constant (GV task; Fig. 2C, blue and green curves). This single unit is strongly modulated by grip type and retinotopic target position.

A second, more complex example neuron (from AIP) is shown in Figure 3. It was grip-type tuned during the movement epoch with a preference for power grip (Fig. 3D–F) and also modulated by spatial factors, especially during the execution of power grips in the CV condition (Fig. 3D). Furthermore, it was spatially modulated during the cue, planning, and movement epochs of various conditions (Fig. 3C, E, F), thus demonstrating, in contrast to the first example, a more complex tuning of spatial representation and in particular the absence of a clear retinotopic coding.

**Population tuning**

To better understand the nature of the tuning for grip type and spatial (target and gaze) position in AIP and F5, we performed a two-way ANOVA \(( p < 0.01)\) with the factors grip type (precision or power grip) and spatial position (13 target and gaze positions combined from CV, TV, and GV) separately for each neuron and task epoch. In general, we found a broad variety of tuning combinations in AIP and F5 (Table 1, Fig. 4A, B). Across all task epochs, 43% of all AIP neurons (151 of 353 cells) were significantly tuned for grip type, whereas 73% (259 of 353 cells) were modulated by spatial (target and gaze) position. In contrast, in F5 more units were tuned for grip type (54%; 316 of 585 cells), whereas only 161 of 585 cells (28%) were modulated by position \(( p < 0.01)\).

These findings persisted when comparing specific task epochs (Fig. 4A, B). In both areas, the number of grip-type tuned cells increased significantly during cue epoch (AIP: 12%; F5: 17%) and strongly peaked during movement execution (AIP: 32%; F5: 46%), whereas the number of position-tuned cells remained rather constant, at least from the cue epoch onward (Fig. 4A, B). In AIP, an average of 45% of all cells were spatially modulated across all task epochs, whereas this number was only ~15% in F5. These findings demonstrate that grip type tuning is a prominent feature of F5, whereas reach target and gaze position tuning is more strongly represented in AIP.

The lower fraction of grip-type-tuned units in AIP compared with our previous study (Baumann et al., 2009) could be due to the different task design. Neurons tuned exclusively for gaze or target position, which comprised a substantial fraction of our dataset, could have been easily missed previously, when spatial factors were not varied. Similarly, the F5 population had a lower fraction of grip type tuned units during the cue and planning...
epoch, but not during movement execution, as reported previously (Fluet et al., 2010).

Furthermore, a substantial fraction of cells in both areas was simultaneously modulated by grip type and spatial position (Fig. 4A, B, horizontal lines), indicating that grip type and spatial tuning are not separated, but often co-represented (intermingled) in the cells of AIP and F5. In addition, the prevalence of spatial tuning did not depend on the particular grip type (data not shown).

Similar results were found for the distribution of tuning classifications (Table 1). A large fraction of AIP cells stayed tuned for position during the whole task (19%) or from cue onward (10%), whereas 9% of the units were tuned exclusively in the movement period (Table 1). In contrast, by far the biggest fraction of grip-type-tuned units in AIP was tuned only during movement execution (23%). In F5, position-tuned neurons were rather rare, with cells tuned throughout the task (4%) or only in the movement epoch (6%) being the biggest fractions. In contrast, grip-type-tuned cells were tuned mainly from cue onward (10%) or exclusively during movement execution (27%; Table 1).

These findings were confirmed when the two-way ANOVA was applied in a sliding window analysis to highlight the tuning onset of individual cells (Fig. 4C, D; see Materials and Methods). Again, in AIP, the dominant effect was tuning for position and most position-tuned cells became selective during the fixation and cue epoch. Fewer cells were tuned for grip type and their tuning onset occurred predominantly during movement execution (Fig. 4C). In the F5 population, we found similar distributions for the onset of position and grip type tuning: position tuning started predominantly in the fixation and cue epoch, whereas cells became grip type tuned most prominently during movement execution and less so during cue presentation (Fig. 4D). Furthermore, in both areas, we found some cells that were position tuned already before the fixation epoch. This could have been due to motor noise before trial onset that the animal might have localized while the target was repositioned (see Materials and Methods).

We found grip-type- and position-tuned cells in AIP and F5. Position tuning was present in all task epochs at approximately constant levels, whereas the number of grip-type-tuned cells was strongly increasing during the task with a peak at movement execution. Position information was more prominent in AIP than in F5, whereas F5 contained a much higher fraction of grip-type-tuned cells compared with AIP. These findings demonstrate that both areas contribute differently to the frontoparietal grasp network (Rizzolatti and Luppino, 2001; Baumann et al., 2009; Fluet et al., 2010).
To estimate how well individual neurons represent grip type and spatial positions, we computed for each neuron a ROC score by comparing spiking activity of the preferred versus nonpreferred grip type (across all tasks) and of the preferred versus nonpreferred target/gaze position (separately for the CV, TV, and GV tasks). Figure 4E,F shows the averaged ROC score of all tuned neurons for grip type as well as for the CV, TV, and GV spatial positions (sliding window analysis). In AIP, the ROC score for grip type showed a small increase after cue onset and a stronger one around movement execution. In contrast, ROC values for all three position tasks (CV, TV, and GV) increased earlier and stronger compared with grip type and stayed at that high level throughout the task (Fig. 4E,F). In F5, ROC scores had approximately the same temporal profile as in AIP, but were higher for grip type and somewhat lower for spatial positions (Fig. 4F). However, there were two differences. First, ROC scores for target and gaze position declined strongly during movement execution in F5 while remaining high in AIP, indicating a weaker representation of spatial positions in F5 during movement execution. Second, the ROC score was larger for grip type than for spatial positions during movement execution in F5. These ROC results, together with the ANOVA analysis, show that both the stronger representation of grip type in F5 and the more prominent representation of spatial information in AIP hold true not only in terms of the fraction of tuned neurons, but also with respect to their tuning strength. Spatial ROC values for CV, TV, and GV differed significantly from chance in both areas (one-way ANOVA, \( p < 0.01 \)) and were quite similar within each area. These findings are consistent with the notion that AIP and F5 might encode relative spatial positions, such as the retinotopic target position (Fig. 2), in addition to spatial and other, more complex gaze-dependent representations (Fig. 3).

![Figure 3](image-url)

**Figure 3.** Example unit from AIP. This neuron is modulated by both grip type and spatial factors. A–C, Neural activity for performing precision grip trials for CV, TV, and GV conditions. D–F, Neural activity for performing power grip trials for CV, TV, and GV conditions. Figure conventions identical to Figure 2.

**Table 1. Cell classification by tuning in task epoch**

<table>
<thead>
<tr>
<th>Fixation</th>
<th>Cue</th>
<th>Planning</th>
<th>Movement</th>
<th>AIP (%)</th>
<th>F5 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>0.4</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>3.7</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>22.9</td>
<td>27.4</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>0.8</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>3.1</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>2.3</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>3.1</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>57.2</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Cell classes are listed according to the presence (+) or absence (−) of significant tuning for GT or position (Pos) in the task epochs: fixation, cue, planning, and movement (two-way ANOVA, \( p < 0.01 \); see Materials and Methods). Percentages indicate the fractional size in the AIP (\( n = 353 \)) and F5 (\( n = 585 \)) population, respectively; percentages add up to 100% within each column.

**ROC analysis**

To estimate how well individual neurons represent grip type and spatial positions, we computed for each neuron a ROC score by comparing spiking activity of the preferred versus nonpreferred grip type (across all tasks) and of the preferred versus nonpreferred target/gaze position (separately for the CV, TV, and GV tasks). Figure 4E,F shows the averaged ROC score of all tuned neurons for grip type as well as for the CV, TV, and GV spatial positions (sliding window analysis). In AIP, the ROC score for grip type showed a small increase after cue onset and a stronger one around movement execution. In contrast, ROC values for all three position tasks (CV, TV, and GV) increased earlier and stronger compared with grip type and stayed at that high level throughout the task (Fig. 4E,F). In F5, ROC scores had approximately the same temporal profile as in AIP, but were higher for grip type and somewhat lower for spatial positions (Fig. 4F). However, there were two differences. First, ROC scores for target and gaze position declined strongly during movement execution in F5 while remaining high in AIP, indicating a weaker representation of spatial positions in F5 during movement execution. Second, the ROC score was larger for grip type than for spatial positions during movement execution in F5. These ROC results, together with the ANOVA analysis, show that both the stronger representation of grip type in F5 and the more prominent representation of spatial information in AIP hold true not only in terms of the fraction of tuned neurons, but also with respect to their tuning strength. Spatial ROC values for CV, TV, and GV differed significantly from chance in both areas (one-way ANOVA, \( p < 0.01 \)) and were quite similar within each area. These findings are consistent with the notion that AIP and F5 might encode relative spatial positions, such as the retinotopic target position (Fig. 2), in addition to spatial and other, more complex gaze-dependent representations (Fig. 3).
Coordinate frames linear modeling

To further investigate the tuning of individual neurons, we modeled the firing rates of each neuron (in specific task epochs) in a stepwise linear model including the factors grip type (GT), target position (T), and gaze position (G) (Fig. 5A; see Materials and Methods). In both areas, a large fraction of cells was significantly modulated by at least one of these factors. In AIP, an average of 68% of cells were modulated across epochs [fixation: 214 of 353 neurons (61%), cue: 239 (68%), plan: 238 (67%), movement: 275 (78%)], whereas in F5, an average of 48% of all cells were modulated by at least one factor [fixation: 181 of 585 neurons (31%), cue: 266 (45%), plan: 287 (49%), movement: 390 (67%)]. Following these results, we categorized each neuron according to its significant modulation by GT, T, G, or T&G.

Figure 5B, C depicts the fraction of cells modulated by these factors for the various epochs. Consistent with our previous analysis (ANOVA, ROC), the fraction of cells with a significant factor GT strongly increased from the fixation (<5%) to the movement epoch (AIP: 40%; F5: 58%), whereas the fraction of cells with significant spatial modulation (T, G, or T&G) was approximately constant throughout the task, average across all spatial groups: ~60% for AIP (fixation: 59%, cue: 63%, plan: 61%, movement: 68%), ~30% for F5 (fixation: 29%, cue: 32%, plan: 31%, movement: 32%).

Concerning the underlying coordinate frame for the coding of target and gaze position, similar fractions of cells were classified as modulated by T, G, or T&G. Across epochs, values averaged ~21% for AIP (T: 15%, G: 26%, T&G: 22%; Fig. 5B) and 10% for F5 (T: 11%, G: 12%, T&G: 9%; Fig. 5C).

Spatial position and grip type modulation were also found to be intermingled: in AIP, spatially tuned groups (T, G, and T&G) exhibited a similar percentage of grip type tuning than was found in the

---

Figure 4. Population results. A–D, Spatial and grip type tuning in the recorded populations in AIP (n = 353) and F5 (n = 585). A, B, Fraction of cells with tuning for the factors grip type (blue) or position (red) during the different task epochs (two-way ANOVA, p < 0.01); short black lines indicate fraction of cells with tuning for both grip type and position. C, D, Sliding window analysis (window size 200 ms, step size 50 ms) for each neuron (y-axis) and time-step (x-axis) revealed the times with significant tuning. Horizontal bars indicate time windows with significant tuning for grip type (blue) or position (red) and are threefold aligned to fixation onset, cue offset, and the “go” cue. Neurons are ordered by tuning onset (defined by the appearance of five consecutive significant steps). Vertical lines indicate onset of fixation, cue, planning, and movement. E, F, Averaged ROCs for the tuned units of A and B. Colored lines show the ROC score of sliding windows analysis for distinguishing the two grip types (blue) and the preferred and nonpreferred condition in the CV task (yellow), TV task (orange), and GV task (red). Curves are threefold aligned to fixation onset, cue offset, and the “go” cue (indicated by black triangles); gaps in the curves mark realignment.
entire population (Fig. 5D). In F5, the same was true for the T- and G- modulated cells (Fig. 5E), whereas the T&G group represented grip type even more frequently (cue: 60%, plan: 64%, movement: 83%) than in the entire population (cue: 26%, plan: 32%, movement: 58%). This again suggests that grip type and spatial information are processed in a combined fashion in AIP and F5.

To further analyze the modulation of the spatially tuned neurons with respect to target and gaze position and their underlying reference frame, we compared the estimated coefficient vectors for target and gaze modulation. For each neuron, we computed an LC index comparing the length of both gradient vectors and the angular orientation difference between these vectors. Figure 5F, G shows a scatter plot of these measures for data during the planning epoch. Purely target- or gaze-coding neurons had a LC close to 0 or +1, respectively. Neurons were considered to be retinotopic (i.e., coding the target position in retinal coordinates) if their target and gaze coefficient vectors had approximately equal length (LC close to 0.33) and opposite directions (angular difference larger than 135°), as indicated by the black rectangle in Figure 5F, G. Although angular difference was broadly distributed for target-modulated (green dots) and gaze-modulated (light blue dots) neurons, a majority of T&G neurons (red dots) was retinotopic (distribution shown for the planning epoch; AIP: 65%, F5: 69%). The percentage of retinotopic neurons was also quite similar in the other task epochs, as reported in Figure 5B, C, dashed red line. Across all task epochs, retinotopic neurons comprised approximately 68% (AIP) and 73% (F5) of all T&G cells, which corresponds to a fraction of 15% (AIP) and 6% (F5) of the entire population (Fig. 5, B, C, dashed line), approximately equal in size to the fraction of target-tuned cells (green line). Furthermore, the amount of grip type tuning in retinotopic neurons was similar to that of the

---

**Figure 5.** Linear model. **A**, Combined linear model fitted to the data. The fit predicted the firing rate of an individual neuron by the factors GT (dark blue underscore), T (green), and G (light blue). **B**, Percentage of cells in AIP that are tuned for each of the factors or for a combination of the two spatial factors (solid red line). Their retinotopic fraction is shown as dashed red line. **C**, Same analysis for F5. **D**, Fractions of grip-type-tuned cells in the spatially tuned populations (T, G, T&G) and in the entire population. **E**, Same analysis for F5. **F**, Scatter plot of the spatially tuned neurons in AIP illustrating angular orientation difference (y-axis) between the target position (t) and gaze position (g) vectors against the LC of these vectors (x-axis). Neurons with target and gaze tuning were considered retinotopic if the coefficient vectors (t and g) were of comparable length (LC < 0.33) and oriented in nearly opposite directions (angular difference > 135°). The fraction of neurons meeting these criteria (inside black rectangle) is drawn as a dashed line in **B** and **D**. **G**, Same analysis for F5.
T&G group in AIP and F5 (Fig. 5D,E). These results clearly demonstrate a large variety of target, gaze, and retinotopic target coding that is strongly intermingled with grip type coding in AIP and F5.

Decoding simulation
To further investigate the coding scheme and the employed coordinate frames in AIP and F5, we performed an offline decoding analysis to predict grip type, target location, and gaze position from our population of sequentially recorded neurons (see Materials and Methods). In agreement with our findings above, decoding performance for grip type was in both animals >75% correct for AIP and nearly perfect (>95%) for F5 (Fig. 6A). For position decoding (Fig. 6B), we found in AIP average performances (across cue, planning, and movement epochs) of 89% (animal P) and 80% (animal S) (13 conditions; chance level: 7.7%), whereas for F5, these values were considerably lower (animal P: 48%; animal S: 38%). Figure 6C illustrates a confusion matrix for decoding the 13 spatial conditions (AIP activity from animal P during the planning period; Fig. 6B, asterisk). The reported overall performance of 92% is the mean of all values in the matrix diagonal.

To further investigate the underlying coordinate frame for the coding of the target position, we also decoded separately the spatial and the retinotopic target position and the gaze position (Fig. 6D–F). In AIP, we found the best decoding performance for predicting the retinotopic target position in all of the epochs cue, planning, and movement (average performance across these epochs: animal P, 92%; animal S, 95%; Fig. 6E,G). In contrast, decoding performance in AIP was markedly lower for predicting the spatial target position (P: 85%; S: 74%; Fig. 6D,G) or gaze position (P: 88%; S: 79%; Fig. 6F,G). Similar results were obtained from area F5, however, with an overall lower performance (Fig. 6D–G). In both areas and animals, decoding of retinotopic target position was significantly better than for target or gaze position, as summarized in Figure 6E (one-way ANOVA, and post hoc t test; p < 0.01).

These decoding results demonstrate that AIP and F5 both represent the grip type and the reach target and gaze position during the different epochs of the delayed grasping task, and that target position is encoded in both spatial and retinotopic representations.

Discussion
We analyzed neural activity in two hand-grasping areas, AIP and F5, during a reach-to-grasp task with systematic variation of grip type and of target and gaze position in extrinsic space. In addition to grip type, the spatial factors target and gaze position were strongly encoded in both areas (Fig. 2, Fig. 3). Although the number of grip-type-tuned neurons increased during the task, the fraction of cells with spatial tuning stayed approximately constant (Fig. 4A–D). Similarly, ROC analysis revealed a stronger encoding of spatial positions in AIP and F5 than of grip type, except for F5 during movement execution (Fig. 4E,F). Further analysis of these surprisingly strongly encoded spatial signals revealed that individual neurons represent them in various reference frames, ranging from purely spatial encodings of gaze position to prominent representations of target position in spatial and gaze-centered (retinotopic) coordinates (Fig. 2). Both areas exhibited also mixed (i.e., intermediate or idiosyncratic) representations (Fig. 3), however, the fractions of all of these spatial representations stayed approximately constant throughout the task (Fig. 5). Finally, consistent with the tuning results, decoding simulations resulted in better predictions of grip type from F5 than from AIP activity, whereas spatial conditions could be more accurately decoded from AIP. In both areas, the retinotopic target position was consistently predicted best, or at least as good as the spatial target or gaze position (Fig. 6).
These results demonstrate that AIP and F5 provide a heterogeneous network in which individual neurons represent grip type together with spatial signals, including gaze and retinotopic and spatial target positions. Such a network could integrate gaze and reach target information during grasp movement planning and execution (Jeannard et al., 1995).

**Target and gaze representations**

To our knowledge, substantial representations of reach target and gaze position in AIP have not been described previously. In fact, an influence of spatial target position was denied by Taira et al. (1990), but those investigators tested only a small number of neurons. Approximately half of all AIP neurons contained spatial information in each task epoch, whereas grip type was encoded only by ~30% of all neurons and was strongest during movement execution (Fig. 4A). In addition, the tuning strength in AIP (ROC score) was on average stronger for target and gaze position than for grip type (Fig. 4E). This was surprising, given that AIP is causally linked to the planning and execution of hand-grasping movements (Taira et al., 1990; Gallese et al., 1994).

Differences in wrist orientation when grasping to different target locations were not observed by visual inspection, but could not be completely ruled out. However, small differences in wrist orientation could not explain the prominent encoding of spatial factors (gaze, target, and retinotopic position), which were even stronger than grasp type tuning in AIP. Furthermore, spatial tuning could not be explained by perceived changes in target orientation (e.g., due to nonfoveal retinotopic presentations; Crawford et al., 2011), because these were minimal in our task (less than a few degrees), which is an order of magnitude smaller than that described previously (Baumann et al., 2009; Fluet et al., 2010).

In F5 of the same animals, we found target and gaze representations with similar tuning strength as in AIP (Fig. 4F), however, the fraction of spatially tuned cells was reduced to ~15% (Fig. 4B). Previously, reach target representations have been reported in PMv for a virtual motor task (Schwartz et al., 2004), and Stark et al. (2007) found strong reach and grasp representations in PMv and dorsal premotor cortex (PMd). The stronger reach representation in PMv, compared with our work, could result from different recording locations, because we have investigated more specifically area F5. Gaze-dependent activity has also been reported for F5 in visual response tasks (Gentilucci et al., 1983; Boussaoud et al., 1993).

Interestingly, considerable fractions of neurons represented both spatial and grip type factors (Fig. 4A, B, Fig. 5D, E), which leads to the question of why target and gaze signals are co-represented. Clearly, AIP and F5 are directly and reciprocally connected and participate in the parieto-frontal network for grasp planning and execution (Luppino et al., 1999; Rizzolatì and Luppino, 2001; Borra et al., 2008). Furthermore, intracortical microstimulation in F5 elicited mostly distal upper limb movements (Rizzolatì et al., 1988; Hepp-Reymond et al., 1994; Stark et al., 2007), and chemical inactivation studies of AIP (Gallese et al., 1994) and F5 (Fogassi et al., 2001) resulted in deficits in hand preshaping and grasping without impairment of reaching. This implies that target and gaze position signals in AIP and F5 are perhaps not causally linked to the generation of reach movements. Rather, they might function as a reference signal (e.g., efference copy) for selecting or generating appropriate grasp movements. Because these signals are present well before the start of the movement, they also cannot represent simple sensory feedback (Mountcastle et al., 1975).

**Reference frames**

In both areas, a considerable fraction of neurons was modulated by a combination of the spatial target and gaze position or encoded the target in retinotopic coordinates (Fig. 5). To our knowledge, these findings are novel for AIP. For F5, only a brief report exists describing gaze-dependent reaching signals in a single animal (Mushiake et al., 1997), whereas gaze-dependent activity has been observed in a visual response task (Boussaoud et al., 1993). In contrast, Fogassi et al. (1992) and Gentilucci et al. (1983) reported gaze-independent response fields in F5. Our results of retinotopic and spatial reach representations in F5 (Fig. 4A, B, Fig. 5D, E) extend these studies by demonstrating that reaching and grasping signals are co-represented by the same neurons, suggesting that signals are computationally combined, as was also shown for neurons in area V6A (Fattori et al., 2010).

Retinotopic reach representations have been observed in other posterior parietal areas related to arm reaching, including area V6A (Marzocchi et al., 2008), the parietal reach region (Batista et al., 1999; Buneo et al., 2002; Pesaran et al., 2006; Chang et al., 2009; Chang and Snyder, 2010), and parietal area 5 (Buneo et al., 2002; Bremner and Andersen, 2012), but also in PMd (Boussaoud and Bremner, 1999; Pesaran et al., 2006, 2010; Batista et al., 2007). Most of these studies found reach representations in various reference frames, ranging from spatial to mixed and purely retinocentric coordinates. In addition, gaze-dependent reach target representations have also been observed in human parieto-frontal cortex (Beurze et al., 2010).

Our results therefore support the idea that a set of spatial and gaze-dependent reference frames (including retinotopic ones) can provide a unified framework for spatial information processing that is able to facilitate multisensory integration and coordinate transformation (Andersen and Buneo, 2002; Buneo and Andersen, 2006; Crawford et al., 2011). Nevertheless, retinotopic reference frames can coexist with other, partly overlapping representations and also with other signals not tested here, such as the initial hand position or head position in space (Pesaran et al., 2006; Chang and Snyder, 2010; Bremner and Andersen, 2012).

Finally, our finding that the spatial coding schemes in AIP and F5 are approximately identical, although less represented in F5, is compatible with the hypothesis that reaching information in F5 originates from AIP, as supported by direct anatomical connections (Luppino et al., 1999; Borra et al., 2008). Alternatively, it could originate from PMd (Matelli et al., 1986; Gharbawie et al., 2011). AIP could receive reaching information from parietal reach region, V6A, or parietal area 5 (Borra et al., 2008; Gambardella et al., 2009).

**Motor planning and coordination**

We found that AIP and F5 contain not only signals directly related to their specific output (grasping code), but also reach target- and gaze-related information. These signals are quite strong in AIP and F5, so strong that the spatial and retinotopic reach target and the gaze position can be decoded from them (Fig. 6). They may be used for the coordination of reaching and grasping movements (Stark et al., 2007). The fact that we found reach-related signals in AIP and F5 points to a distributed nature of motor planning, which cannot be clearly separated into a grasping (dorsolateral) and reaching system (dorsomedial network). Similar co-representations of reaching and grasping signals have also been described in the dorsomedial network both in monkeys (parietal area V6A) (Galletti et al., 2003; Fattori et al., 2010; Fattori et al., 2012) and humans (Grol et al., 2007; Verhagen et al., 2008). Inter-areal computation between parietal and premotor...
areas to generate reach-coordinated grasping movements seems unlikely. However, similar coordination seems in place between parietal and premotor cortex to integrate external (e.g., sensory) and internal (e.g., volitional or intentional) signals for reach and grasp movement planning (Rizzolatti and Luppino, 2001; Andersen and Cui, 2009). We conclude that the classical subdivision into reaching and grasping areas is certainly lost from a representational point of view. From a causal perspective, however, future research is needed to investigate the effective connectivity of these areas with respect to grasp, reach, and gaze movements. At least, the sheer presence of a motor signal does not imply a causal effect for movement execution.

References
Lemmhein and Scherberger • Reach and Gaze Signals in Macaque Grasp Areas