Title
Mixed reference frame representations underlie the use of multimodal sensory signals for reaching.

Author
McGuire, Leah Marie Medvick

Publication Date
2010

Peer reviewed|Thesis/dissertation

Mixed reference frame representations underlie the use of multimodal sensory signals for reaching.

by

Leah Marie Medvick McGuire

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Neuroscience

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

[redacted]

Acknowledgements

For support (both emotional and technical) and many helpful discussions:

Zack Chadick, Daniel McGuire, Pat Medvick, Helen Shen, Matt Fellows, Tim Verstynen, Lauren Hoffman, Maria Simani, Audrey Carstensen, Blakely Magliaro, Charles Biddle-Snead, Maria Dadarlat, Sergei Rebrik, Liz Montgomery, Ken McGary, Lazlo Bocskai, Darrell Floyd

For mentoring and instruction:

Philip Sabes, Steve Lisberger, Loren Frank, Allison Doupe, Larry Snyder

The text of Chapter 1 of this thesis is a reprint/rewrite of the material as it appears in "Sensory transformations and the use of multiple reference frames for reach planning" *Nat. Neurosci.*, 2009. The coauthor listed in this publication directed and supervised the research that forms the basis for the thesis.

Mixed reference frame representations underlie the use of multimodal sensory signals for reaching.

Leah Marie Medvick McGuire

Abstract

The sensory signals that drive movement planning arrive in a variety of 'reference frames', and integrating or comparing them requires sensory transformations. I set out to examine how the different forms of sensory signals, and the transformations needed to compare them, affect the representation and integration of multimodal sensory information. I used a combination of human psychophysics and electrophysiological recordings from rhesus macaques to examine how visual and proprioceptive information are used for reach planning and execution.

The human experiment was designed to exploit stereotyped patterns of gaze-dependent reach errors to determine whether the reference frame representations for reach planning depend on the visual and proprioceptive sensory information available. The results of this experiment were interpreted with a model of reach planning in which the statistical properties of sensory signals and their transformations determine how these signals are used. I found that no single reference frame representation was adequate to explain the observed error patterns when visual and proprioceptive information were

varied. Only by integrating movement plans across multiple reference frame representations was the model able to capture the observed error patterns (Chapter 1).

Taking the results from this model, I next looked for evidence of these multiple reference frame representations across different sensory modalities in sensory-motor cortical areas of the rhesus macaque (Area 5 and MIP). I found that neurons in these areas use mixed reference frame representations, which are consistent across reaches to targets specified by different sensory modalities (visual and/or proprioceptive targets, Chapter 2). Additionally, I found that integration of multimodal sensory signals in Area 5 and MIP emerges primarily across the population response rather than within individual cells' responses (Chapter 3). These findings are consistent with the model results showing that sensory information is integrated in multiple reference frame representations, regardless of the reference frame in which sensory information enters the nervous system. These results illustrate one way that the brain can represent and integrate sensory information arriving from different sensory modalities.

Table of Contents

List of Figures and Tables

Chapter 1: Sensory transformations and the use of multiple reference frames for reach planning.

Supplemental Information for Chapter 1: Sensory transformations and the use of multiple reference frames for reach planning.

Supplemental Information for Chapter 3: Integration of visual and proprioceptive information in population responses of the posterior parietal cortex.

Introduction

When interacting with their environment animals use information from multiple sensory modalities. For example, if a mosquito lands on a persons arm he may have somatosensory information about the location of the mosquito on his arm, which can be used to quickly swat at the mosquito. Similarly, if he sees the mosquito land on his shirt sleeve he can use this visual information to perform the same movement. If the person both feels the mosquito on his arm and sees the mosquito he can use both pieces of sensory information to plan and execute a more accurate movement, increasing his chances of squashing to mosquito. All of these scenarios illustrate how very different sensory stimuli, specifically vision and somatosensation, can be used to plan the same movement.

Despite the ease with which visual and somatosensory, or more specifically proprioceptive, information are used to plan and execute movements in everyday life, these sensory inputs enter the nervous system in very different forms, and little is known about how they are represented in the motor circuit. The information arriving in different sensory modalities comes from different types of sensory receptors, e.g. photoreceptors in the retina for visual information and mechanoreceptors in the muscles and joints for proprioceptive information. In addition to signaling different types of stimuli these receptors signal different types of spatial information. Visual information enters the nervous system as a cite of stimulation on the retina and so conveys information about location relative to the position of the eyes, i.e. in an eye-centered or retinotopic reference frame, while proprioceptive information enters the nervous system as information about

joint angles, which can be interpreted to give positional information relative to the torso, i.e. a body-centered reference frame. The spatial information from these sensory modalities follows discrete pathways to the primary sensory areas, the visual and somatosensory cortices, respectively, where visual and proprioceptive information are first represented in the cortex [1]. Both of these areas have inputs to the posterior parietal cortex (PPC), which is thought to integrate visual and proprioceptive information for movement planning [2-7]. Given the different types of spatial information contained in these sensory inputs, it is unclear what reference frame they should be represented in when they converge in the PPC. The focus of my thesis has been to determine how visual and proprioceptive information are represented in the PPC for use in reach planning and execution.

There is conflicting evidence about whether visual and proprioceptive information are represented in different reference frames (each reflecting the reference frame of the sensory input) or the same reference frame (which could be eye-centered, body-centered or a mixture of reference frames) for movement planning. While the majority of studies of reference frames in the PPC have focused on visually guided movements [8-14] a few studies have compared the reference frame representations for movements to visual and auditory targets [15-17]. These studies found, alternately, that a common reference frame was used for both visual and auditory targets [15] or that reference frames partially shifted to resemble the reference frame of visual or auditory information [16, 17]. A partial shift in reference frame is possible because representations are often intermediate between pure reference frame representations, i.e. pure eye-, hand-, or body- centered coding [4, 11-13, 16, 18-

[27]. When responses depend partially on eye, hand, or body position, the importance of each of these variables in determining responses can vary, resulting in a partial reference frame shift, with different sensory inputs. Representations that partially shift toward the reference frame of the sensory input being encoded are common in multimodal sensory areas [16, 17, 27-29]. This suggests that common reference frame representations are not necessary for multimodal representations of spatial information [27]. Further, psychophysical studies of reaches to visual and proprioceptive targets suggest that different reference frame representations may be used for these tasks [30, 31]. Thus, I first asked which reference frames are used during reaches with different amounts of visual and proprioceptive information.

In Chapter 1, I looked at how the reference frames used for reach planning depend on the sensory information available with a human psychophysics experiment. Changes in eye-dependent reach errors were used to elucidate whether reference frames depend on the visual and proprioceptive information available for movement planning. This data was interpreted with a model of reach planning in which sensory information is represented in multiple possible reference frames. The pattern of eye-dependent errors observed in this experiment can only be explained by the model when the same set of reference frames are used to integrate sensory information and plan reaches, independent of the sensory information available. Further, the differential weighting of sensory information observed previous psychophysical studies [31] emerges naturally from this model because of the differential reliability of reference frame representations when different sensory information is available. These results suggest that the same reference

frames should be used to represent information for reach planning, regardless of the sensory information specifying the movement.

In Chapter 2, I set out to directly measure the reference frames used by single cells in the PPC of rhesus macaques during reaches to targets specified by visual and proprioceptive information. Studies of the reference frame representations used for visual reaches in the superior parietal lobule (SPL) suggest that areas in the intraparietal sulcus, such as V6A and MIP, tend to have reference frame representations with more dependence on eye position [8, 11, 23, 32] while the surface of the SPL, Area 5, tends to have reference frame representations with more dependence on hand or body position [10, 12, 13]. Thus, Area 5 and MIP appear to use different reference frame representations, and comparing responses to visual and proprioceptive target reaches across them allowed me to distinguish between two possible representation schemes. In the first scheme, the same reference frames are used regardless of sensory information, as predicted by the model in Chapter 1. In the second scheme, representations are flexible and depend on the sensory information available, as described in other multimodal areas [16, 17, 27-29]. I found a mixture of reference frames in both Area 5 and MIP, but the composition of this mixture was independent of whether reaches were to visual or proprioceptive targets. Further, tuning curves were largely aligned across target modalities. Thus the first representation scheme seems to describe responses in the SPL, reference frames do not depend on the sensory information used to specify reach target.

One explanation for these shared reference frames across target modalities is that they would simplify the integration of sensory information from multiple modalities [33]. In Chapter 3, I looked at whether sensory integration was occurring in Area 5 and MIP when multiple sources of sensory information were available. Individual cells displayed a range of changes in responses to bimodal versus unimodal sensory information in both of these areas. However, it was the mixture of modality preferences in the population that resulted in the greatest increase in population responses for bimodal targets. This increase in the population response could provide a stronger position signal for downstream motor areas when planning movements, suggesting that the shared representations in Area 5 and MIP may be important for integrating sensory information.

The results of these studies support the conclusion that visual and proprioceptive information are represented in multiple common reference frames for integration of sensory information and movement planning. However, rather than being a series of discrete reference frame representations, as implemented in the simple model of reach planning (Chapter 1), these reference frames take the form of a continuum of intermediate reference frame representations (Chapter 2). Transforming both visual and proprioceptive inputs into these mixed representations may aid in integrating sensory information from these two sensory modalities (Chapter 3) and in maintaining unified visual and proprioceptive percepts about the spatial locations of the hands [30, 31, 34-36]. Thus, despite entering the nervous system in very different forms, visual and proprioceptive information appear to converge in multiple reference frame representations in the PPC for movement planning and execution.

Chapter 1:

Sensory transformations and the use of multiple reference frames for reach planning.

Introduction

Humans use a variety of sensory signals when interacting with the environment. We can just as easily reach to pick up a coin we see in front of us as we can transfer the coin from one hand to another without looking. Using multiple sensory modalities for planning similar movements is problematic, since different sensory signals arrive in different "reference frames". Specifically, early visual pathways represent stimulus location relative to current gaze location – a *retinotopic* representation, while proprioceptive signals represent hand location relative to the shoulder or trunk – a *body-centered* representation. In order to combine or compare such signals, some of them must be *transformed* between reference frames. While such transformations may appear to be mathematically trivial, here we show that transformations can incur a cost by adding bias and variability [31, 37] into the transformed signal. Thus, such transformations likely play a major role in determining the flow of information in motor planning circuits.

It has been argued that transforming sensory signals into a common representation would simplify reach planning [33, 38-41], and many researchers have attempted to

characterize this representation. Psychophysical studies addressing this issue have assumed that the pattern of reach errors reflects the reference frame of the neural representations for reach planning. However, these studies have argued for both retinotopic [42-46] and hand- or body-centered [47, 48] planning. Primate physiology and human fMRI studies have also found evidence for a range of neural representations for movement planning [8, 10, 13, 24-27, 32]. This abundance of representations makes it seem unlikely that reach planning uses a single common representation [19, 20, 22, 25, 49]. We provide a quantitative explanation of how noisy sensory transformations make it advantageous to have the same movement plan represented in multiple reference frames. Furthermore, we show that the best representation of movement plan depends on the availability and reliability of sensory signals. We test this idea with an experiment in which different sensory signals specify the same desired reaching movement.

The focus of our experiment is on how gaze-dependent error patterns change with available sensory information. A well-studied gaze-dependent error is the *retinal eccentricity effect*, where subjects overestimate the distance between the center of gaze and a visually peripheral target when pointing to the target [50, 51]. These errors are most parsimoniously described in a retinotopic representation, and they have been cited as evidence for retinotopic reach planning [42, 44, 45, 52]. This is just one example of the commonly made argument that the reference frame of a movement plan directly determines the spatial pattern of errors [40, 46-49, 53-56]. We found that the magnitude of these gaze-dependent errors decreases when multiple sources of sensory information about hand [42] or target are available and that the direction of the errors changes when reaching

to proprioceptive targets. These results are interpreted with a model of movement planning where the available sensory signals are combined in a statistically principled manner in two separate reference frames. The model provides a novel explanation for these gaze-dependent reach errors: they arise when transforming sensory information about target location between representations due to a biased internal estimate of gaze direction toward the target. These findings demonstrate that spatial pattern of reach errors does not necessarily reflect the reference frame of the underlying neural representation.

Results

Measuring gaze-dependent reach errors

We first examined how the pattern of reach errors depends on the sensory signals available during the planning and execution of a movement. Specifically, we manipulated information about target location and initial hand position, the two variables needed to compute a movement vector. Target information was varied by having subjects reach either to visual targets (VIS), proprioceptive targets (the index finger of the left hand, PROP), or targets consisting of both visual and proprioceptive signals (the left index finger with visual feedback, VIS+PROP). Information about initial hand position was varied by having subjects reach either with (FB) or without (NoFB) visual feedback of the right (reaching) hand before movement onset, although feedback was never

available during the movement. For each of the six resulting trial types (Fig. 1b), we measured movement errors as subjects reached to an array of targets with gaze held on one of two fixation points (Fig. 1a).

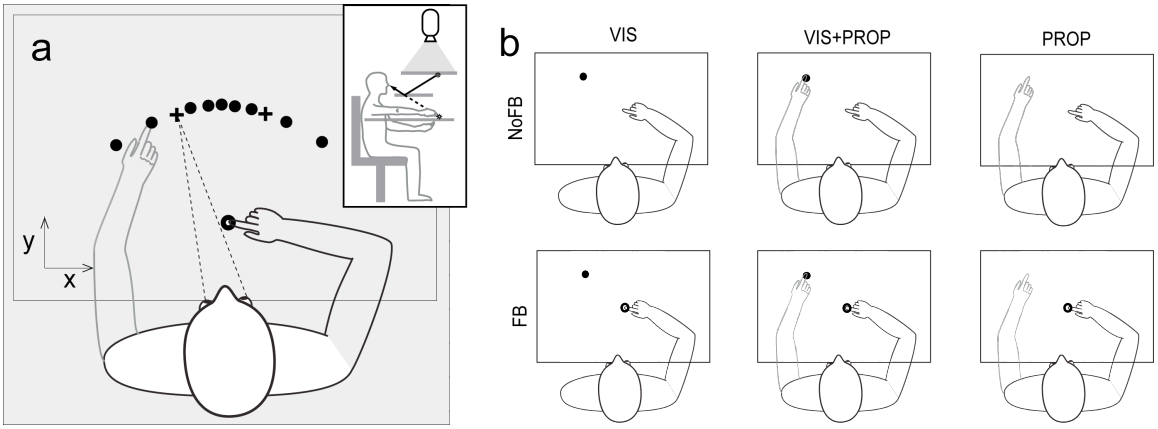


Figure 1: Experimental setup. a) Array of eye and hand targets; + fixation points, • reach targets, o start target. Inset: schematic side view of subject in experimental rig, right arm resting on table, left hand below the table. b) Six task conditions; • visual targets, o visual feedback about position of left hand at start position.

A comparison of reach endpoints at the midline target for an example subject illustrates that reach errors depend on the sensory signals available for movement planning (Fig. 2). The errors differ markedly between gaze locations, and these gaze-dependent effects change across trial types. When a visual target is available (Fig. 2a-d), reach endpoints are biased *away* from the gaze location, i.e., the retinal eccentricity effect described above, and the magnitude of the effect decreases with increasing sensory information (Fig. 2a vs. 2b-d). When only a proprioceptive target is available, this effect is not seen, and, if anything, a small bias in the opposite direction is observed (Fig. 2e-f). These differences are consistent across targets (Supplemental Fig. S1). In addition to these gaze-dependent effects, there is a gaze-independent bias in reaching that may differ across targets and trial types. This effect can be observed either by averaging the endpoints in the two fixation conditions or with a separate gaze-free trial condition

(Supplemental Figs. S1, S2). While there is a trend toward overshooting the target, the pattern of this bias, across targets and trial types, is idiosyncratic from subject to subject (Supplemental Fig. S2), making these patterns difficult to interpret. We therefore focus on the more consistent gaze-dependent effects.

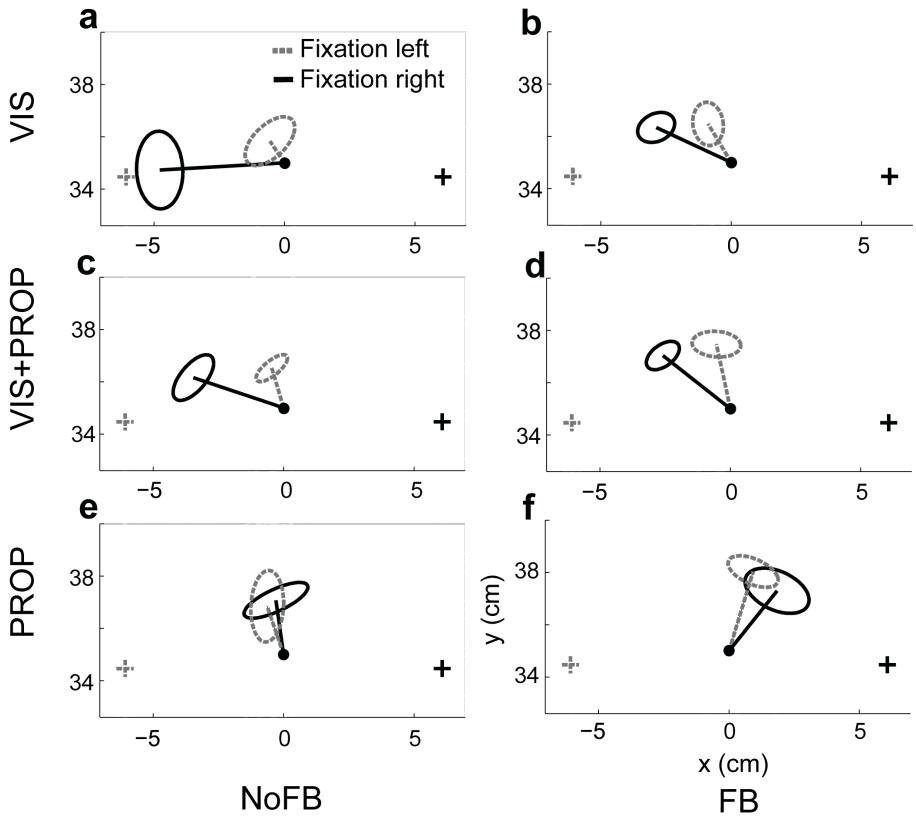


Figure 2: Reach errors at center target for an example subject for all trial types. Lines indicate mean reach error for each gaze-position, ellipses represent standard deviation, + fixation points, • reach targets (solid black lines, gaze right; dashed grey lines, gaze left). The origin (not shown) is located directly below midpoint of the eyes.

In order to isolate the gaze-dependent effects, we transformed the reach errors into polar coordinates about the midpoint of the eyes and subtracted the gaze-free errors (*Methods* and Supplemental Fig. S3). The resulting angular reach errors are significantly different between gaze locations for all trial types, though the pattern varies across trial types (Fig. 3). In VIS/NoFB trials (Fig. 3a), the retinal eccentricity effect is clearly seen in the angular errors: subjects made rightward (positive) reaching errors when fixating to the left of the target and leftward (negative) reach errors when fixating to the right of the target. In VIS/FB trials, the errors for the two gaze locations follow a similar pattern, but

with smaller magnitude. (Fig. 3b, and [42]). In contrast, when subjects reached to exclusively proprioceptive targets (PROP trials) the reach endpoint is *closer* to the fixation point, rather than further from it (Fig. 3e-f). The gaze-dependent error patterns in the VIS+PROP trial types appear to be a combination of those observed in VIS and PROP trials (Fig. 3c-d). In addition to these gaze-dependent effects, there is a general leftward bias in reach error observed in all trial types. This effect is discussed in Supplemental Section 1.2. We also examined errors in depth, i.e. the radial component of reach errors. While there is an overall trend to overshoot the target, these errors do not differ between the two fixation conditions (Supplemental Fig. S4). In other words, the gaze-dependent errors are confined to the angular component of the reach error.

For all trial types, the gaze-dependent error patterns qualitatively align when angular error is plotted in a retinotopic reference frame, i.e. as a function of target relative to gaze (Fig. 3, insets). The fact that these errors look like a function of the retinal eccentricity of the target appears to support a retinotopic representation for reach planning [42, 44, 45, 52]. However, since these patterns differ markedly across trial types they cannot be readily explained in terms of a bias arising solely from a retinotopic representation. This suggests the need for another explanation of these apparently retinotopic errors.

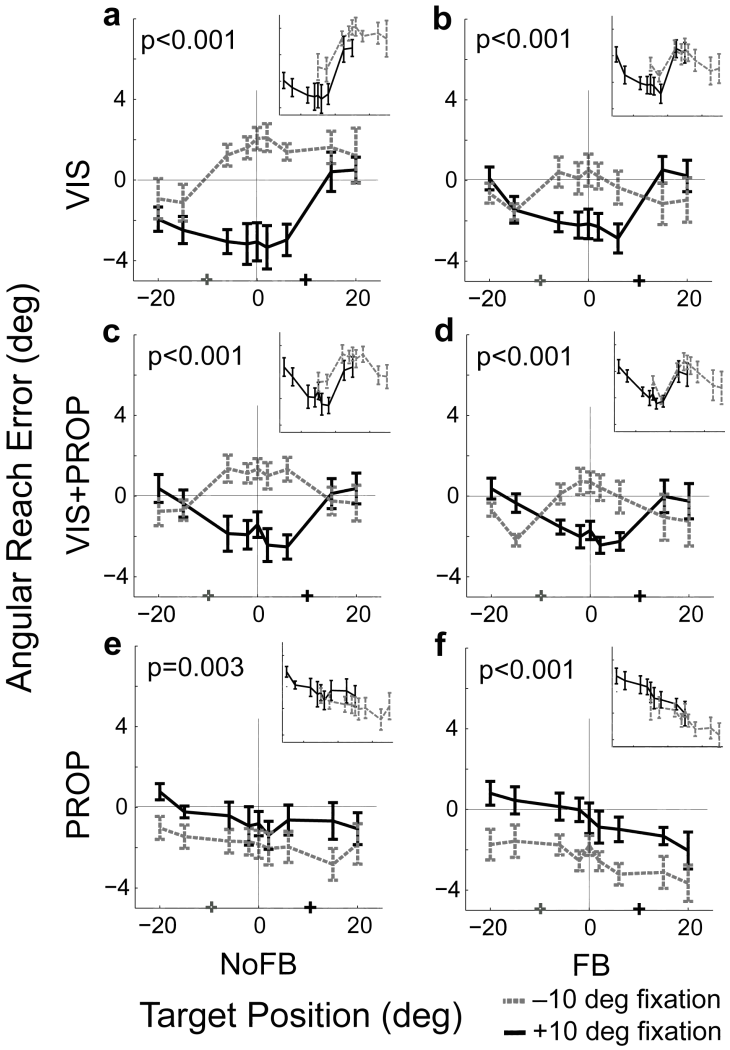


Figure 3: Average angular reach error across subjects for each trial condition. Negative values indicate reach endpoints to the left of target, positive values indicate reach endpoints to the right of target. Before averaging, interpolated gaze-free errors for each trial type were subtracted. Error bars indicate standard errors. p-values determined with a paired permutation test.

A Planning Model: Optimal Integration Across Multiple Reference Frames

Here we present a model of reach planning that accounts for the pattern of gaze-dependent errors observed in our data. The model has two key features: the presence of multiple representations for movement planning and a bias in the transformation between those representations.

The model begins with sensory inputs signaling target location, the initial hand position, and gaze location (Fig. 4a). As sensory signals are inherently variable, we model the information they carry as a Gaussian likelihood of true location given the sensory input, with variance reflecting the reliability of that sensory modality (for review see [57]). Visual signals arrive in a retinotopic representation and proprioceptive signals arrive in a body-centered representation. Each available signal is also transformed into the "non-native" reference frame (Fig. 4b). Since subjects' head positions are fixed during the experiment, this complex nonlinear transformation [58] can be approximated by adding or subtracting the gaze location, i.e. by convolving their distributions (Fig. 4b, [37] and *Methods*). When both sensory modalities are available, the "native" and transformed signals are integrated in both representations (Fig. 4b). Movement vectors are then computed within each representation (Fig. 4a). For each of these computations, we assume independence between the signals being combined (see Supplemental Sections 2.4 and 2.5 for discussion).

Since the transformation between retinotopic and body-centered representations relies on an uncertain estimate of gaze direction (Fig. 4e,d), this transformation adds variability to the transformed signal (Fig. 4b). Because of this cost, the location of any given stimulus is more reliably represented in one or the other of these reference frames, depending on the availability and reliability of visual and proprioceptive signals related to the stimulus (Supplementary Equations 1-2). In line with this idea, the present model quantifies how the reliability of a movement vector plan, in either of these representations, depends on the available sensory inputs.

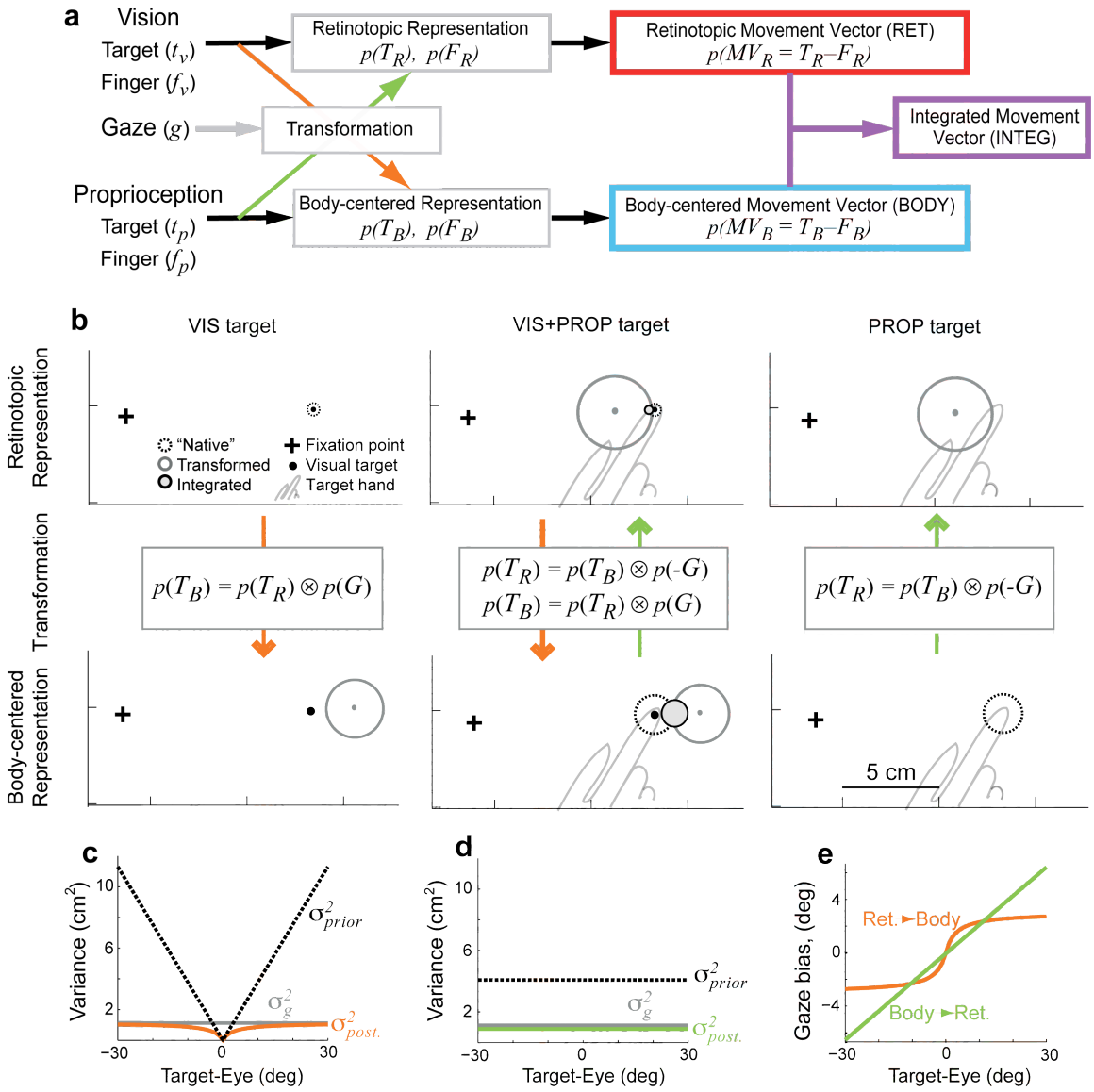


Figure 4: Movement vector planning in multiple coordinate frames. a) Shows flow of information for movement planning. From left to right: Sensory information is represented in native coordinate frame (black arrows) and transformed into non-native representation (color arrows). Combined estimates are used to compute a movement vector in each coordinate frame. Final movement plan may be read out from a single representation (RET, BODY) or using both representations (INTEG). b) Bias and variance introduced during transformation of target information for all target conditions with the INTEG model readout. When multiple sources of target information are available (VIS+PROP trials) information is integrated to form a combined estimate of location. Subjects would be seated 35 cm from the target and fixation point shown. c,d) Posterior variance in estimated eye position (i.e., "transformation variance", colored lines), gaze likelihood variance (grey), and variance of gaze prior (dashed) (see *Methods*) for c) transformation from retinotopic to body-centered space, and d) transformation from body-centered to retinotopic space. e) Bias in eye position estimate used in sensory transformations. All values in this figure reflect the INTEG fits.

In addition to adding variance, the transformation can introduce a bias into the estimates of transformed variables. In particular, we posit that the internal estimate of gaze direction used to transform target information is biased toward the target (Fig. 4e). This bias takes the form of a Bayesian prior on gaze location [37] centered on the target, consistent with the observation that during gaze-free reaches subjects typically fixate reach targets [59]. Since this prior estimate relies on knowing where the target is, the variance of the prior, and hence the magnitude of the bias, is assumed to scale with the variance in the internal estimate of the target location being transformed (Fig. 4c,d). We discuss possible origins for this bias more fully below; here we consider how it would affect movement planning. Since the transformation between representations is simply the addition or subtraction of gaze direction, the bias in transformed signals will either be away from or towards the actual target location depending on the direction of the transformation (Fig. 4b-e). Only "non-native" (transformed) target representations will be biased. Thus, the gaze-dependent errors in the model depend on the availability and reliability of sensory inputs and on the method of reading out the final movement vector.

We consider three possible output schemes for reading the planned movement vector from the model (Fig. 4a). The retinotopic (RET) or body-centered (BODY) representations can each be read out directly. Alternatively, the two movement vector estimates can themselves be combined to form an integrated readout (INTEG). The INTEG readout is a maximum likelihood estimate, where the contributions of the two representations to the final movement vector plan depends on their relative reliability (Supplemental Fig. S6).

Each of the three potential readouts provides a quantitative prediction of the reach errors as a function of hand, target, and gaze locations. The only free parameters in the model are the variances of individual sensory inputs and the gaze priors (*Methods*). The values of the proprioceptive variances are based on previously reported values [60]. Four parameters remain: visual variance, gaze variance, and two scaling factors relating the variance of the gaze prior to the variance in visual and proprioceptive target signals. We fit these parameters to the gaze-dependent errors shown in Figure 3 after mean correction (see *Methods*), using least-squares regression (see Supplemental Table-S1, and Supplemental Fig. S5 for fit values).

Model Fits of Constant and Variable Reach Errors

We first consider how well the three output models fit the observed patterns of gaze-dependent error differences (Fig. 5). When the model is fit with a single-representation readout, RET or BODY, it fails to predict errors for all trial types. This is because only transformed target signals contain a gaze-dependent error. Thus, for VIS targets the RET readout has no error (red line, Fig. 5 a-b), and for PROP targets the BODY readout has no error (blue line, Fig. 5 e-f). While both readouts contain transformed target signals in the VIS+PROP trial types (Fig. 5 c-d), the errors in the retinotopic representation would be in the wrong direction (target biased towards gaze). Thus, fitting the RET readout to the data drives the fit visual variance toward zero, which causes the model to effectively ignore the seemingly more variable proprioceptive signals

(Supplemental Table S1, Fig-S5). A model that switches between these two readouts depending on task also seems unlikely. First, both the RET and BODY readouts fail to capture the differences in the magnitude of gaze-dependent errors that we observed between FB and NoFB conditions. Second, in order to predict the observed errors, the switching scheme would need to rely predominantly on the more variable "transformed" signals, rather than "native" signals, a sub-optimal arrangement (see Fig-6 below). In contrast, when both output representations are combined in the INTEG readout, the model performs well for all trial types. This readout captures both changes in the magnitude of gaze-dependent errors (VIS vs. VIS+PROP, and FB vs. NoFB) as well as the sign reversal observed with PROP targets. It accomplishes this by differentially weighting the two reference frames across trial conditions (Supplemental Fig. S6).

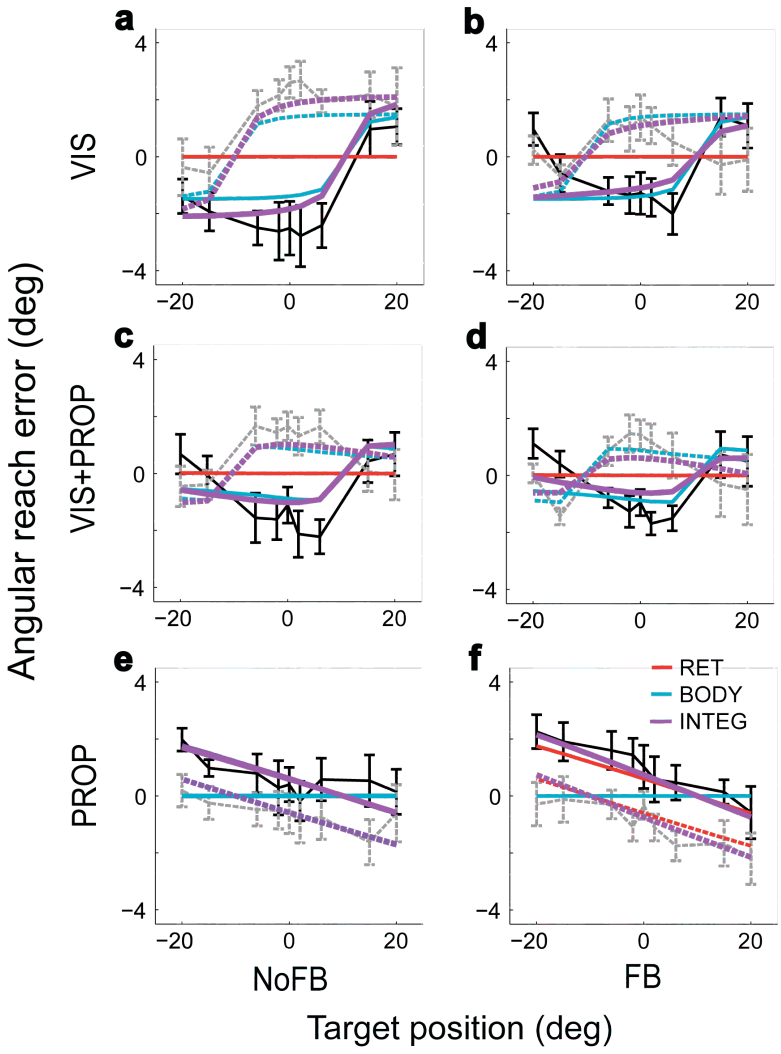


Figure 5: Model fits for gaze-dependent reach errors. Black and gray lines show mean (standard error) errors across subjects for each trial condition, after subtracting the overall mean separately for each of the six trial types. Colored lines show best fit model predictions: RET-red, BODY-green, INTEG-blue. Solid lines, gaze right; dashed lines, gaze left.

In addition to fitting the gaze dependent error patterns, the model predicts the differences in movement variability across trial types (Fig. 6a). Since computations within the model are assumed to be noise-free, model output variability is due entirely to variability in the sensory inputs, shaped by the model computations, and does not require any additional parameter fitting. The INTEG model fit provides an accurate prediction for the changes in output variance across trial types, while the two single-representations fits do not. Of course these predictions come from separate fits (to error) for each readout. The parameters used in the model presumably reflect actual variances in the neuronal representations of sensory inputs. For any given value of these variances, e.g. those obtained with the INTEG fit, we can ask what the variability in the movement plan is for each readout. We found that the INTEG readout generally yields a lower variance estimate (Fig-6b), since it makes better use of all available sensory signals (although the range and magnitude of this advantage depends on the statistical properties of the sensory transformations; Supplemental Fig. S7). In contrast to the idea that a single coordinate frame should dominate movement planning [8, 10, 33, 38, 40-45, 52, 53], this analysis illustrates that having multiple representations of a movement plan would yield more reliable performance across task types.

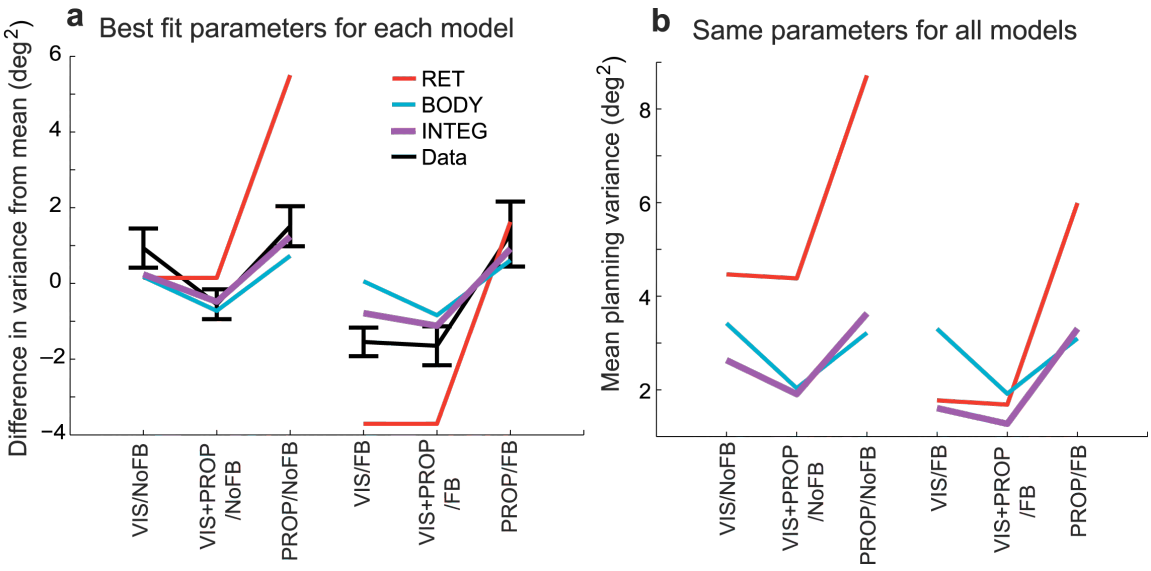


Figure 6: Reach variability. a) Differences in angular reach variance across trial types. For each subject, the overall reach variance was subtracted from the variance for each trial type, and the black lines are the resulting mean (standard error) across subjects. This measure emphasizes task differences in reach variance. Colored lines represent model predictions for variance differences: RET-red, BODY-green, INTEG-blue. b) Mean model planning variability when a single set of variance parameters (INTEG fit) are used for all readouts.

Lastly we tested the model, fit to our own dataset, on an independent dataset from Beurze et al. [42], which contains visual target trials with an expanded range of movements (i.e., more start, and gaze locations). Beurze et al. [42] observed the standard retinal eccentricity effect in their data and showed, as we did above, that the effect magnitude is smaller when visual feedback of the hand is available (data reproduced in Fig. 7a-b). In addition, they found a component of the reach error that correlates with the relative positions of the hand and target (data reproduced Fig. 7c-d). They suggested that these errors could be explained by the introduction of errors in the conversion of target and hand positions into a retinotopic representation of movement planning. As shown in Figure 7e-h our model captures the all of the key features of their dataset, including both the retinotopic and hand-centered error components.

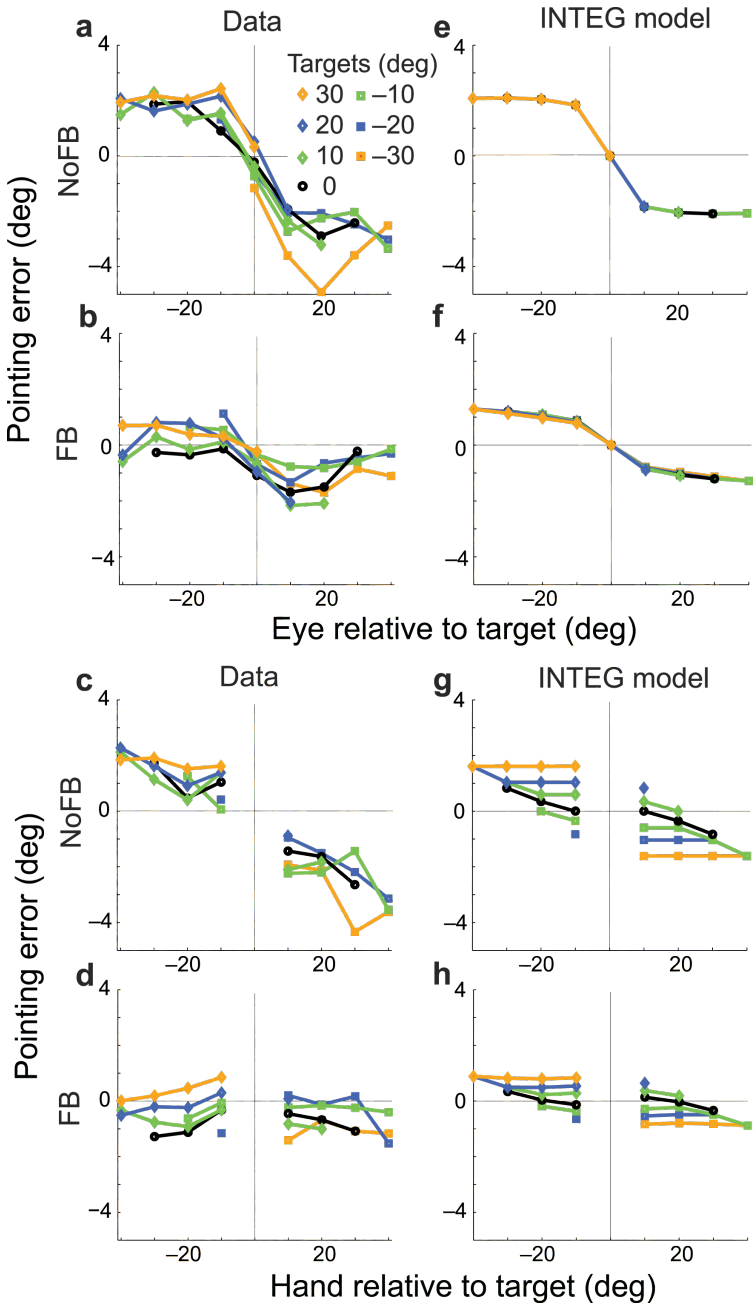


Figure 7: Model predictions for data from Beurze et. al. (2006). a-d) Gaze dependent pointing error. e-h) Pointing error as a function of initial hand position relative to the target. c,d,g,h) average pointing error from Beurze et al. a,b,e,f) INTEG model predictions of pointing error, with parameters fit to our data.

The cost of transformations

Our model of sensory transformations allows us to explain another very different empirical result, again without additional parameter fitting. Sober and Sabes [31] reported

that the relative weighting of visual and proprioceptive feedback of hand location during reach planning depends on the nature of the target. In particular, they showed that vision is weighted more heavily when reaching to visual targets (as in VIS/FB trials here) than when reaching to proprioceptive targets (as in PROP/FB trials here). Sober and Sabes [31] proposed that this difference was due to the cost of performing sensory transformations – the sensory modality that matches the target is weighted more because the other sensory signal has to be transformed. The reach planning model presented here makes this cost explicit: the transformed signal is more variable due to uncertainty in the internal estimate of gaze direction (Schlicht and Schrater [37] made the same argument with a similar planning model). We were therefore able to use this model to make quantitative predictions of the angular error that should result from the artificial shifts in visual feedback used in Sober and Sabes. Figure 8a compares the empirical data and the INTEG model readouts. For both, visual feedback shifts have a weaker effect when reaching to proprioceptive targets than when reaching to visual targets. This effect is quantified in terms of overall weighting of visual versus proprioceptive feedback (Fig. 8b), which is much greater for VIS targets than for PROP targets. In the model, this re-weighting is due to the tradeoff between the retinotopic and body-centered representations of the movement plan (Supplemental Fig. S6), evidenced by the fact that neither the RET nor BODY readout exhibited the effect. This result provides further support for the use of multiple representations in movement planning.

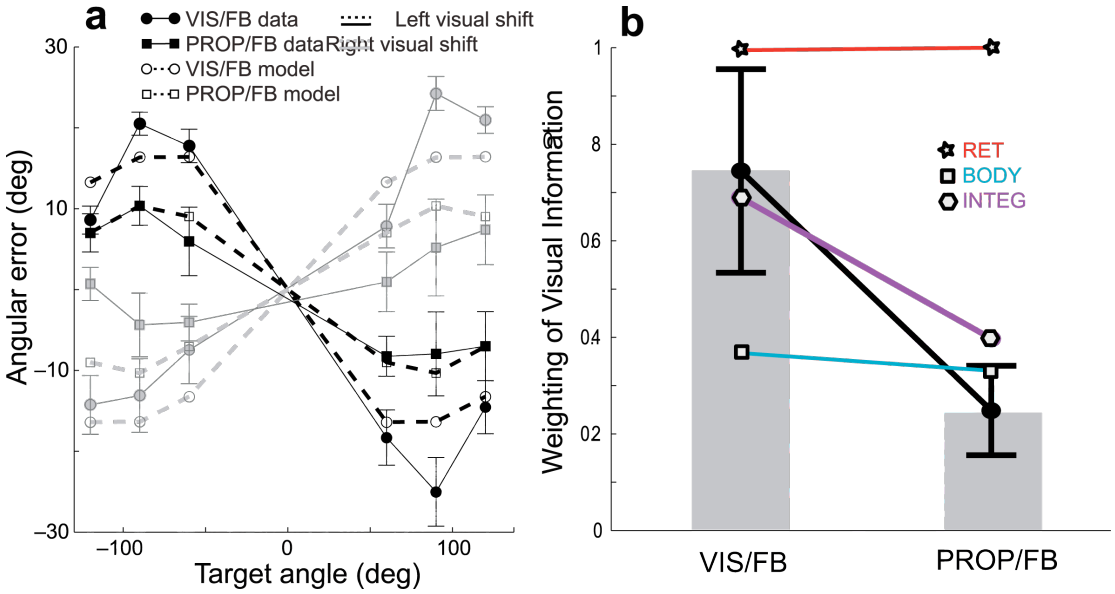


Figure 8: Changes in sensory weighting with target modality. a) Mean angular error induced by artificial shifts in the visual feedback of the hand prior to movement onset. Data from Sober and Sabes (2005). Error bars represent standard errors. Model predictions use the INTEG readout, parameters fit to our data. b) Relative weighting of visual vs. proprioceptive information about initial hand position in movement planning for reaches to VIS and PROP targets. Error bars represent standard deviation across subjects. Colored lines show model predictions for each readout scheme: RET-red, BODY-green, INTEG-blue.

Origins of the Gaze Bias

We have shown that a bias in the internal estimate of gaze location can account for the complex pattern of gaze-dependent reach errors we observed across trial types. We now consider several possible origins of this bias and discuss additional evidence for its presence. This bias might arise due to ether a "covert" saccade plan toward the target or a shift of attention to the target. We tested these hypotheses by controlling the saccade target or the locus of attention independent of the reach target, but these manipulations did not alter the reach error pattern (Supplemental Fig. S8). Alternatively, the bias could arise from a prior expectation that reach targets tend to be foveated, reflecting the fact that eye and hand movements tend to be tightly linked [59, 61, 62]. A Bayesian prior [63] on the

internal estimate of gaze location used for target transformations can account for our observations (see *Methods*). While we were not able to determine the source of the putative gaze bias directly, we were able to observe the bias using an independent measure. Specifically, we found that a visually peripheral reach target biases a subject's estimate of straight ahead, and this bias is consistent with a shift in the estimated gaze direction [64] towards the target (Supplemental Fig. S9). Both this perceptual effect and the constant errors observed in our reach experiment are well modeled by a Bayesian prior on gaze location centered at the target.

Discussion

This study was aimed at testing two widely held ideas in the field of sensorimotor control: that the spatial pattern of errors for a given movement closely reflects the underlying neural representation [42, 44-48, 52, 55], and that there should be a single reference frame for representing movement-related variables for movement planning [8, 10, 33, 38, 40-45, 52-54]. We have argued that neither of these ideas is correct. First, we have shown that a single, apparently retinotopic, pattern of reach errors can be explained by a model in which multiple neural representations are used, e.g. a combination of both retinotopic and body-centered reference frames. Second, we have shown that using more than one representation confers an advantage in terms of planning variability.

Spatial patterns of reach errors have often been cited as evidence that underlying

neural representation are expressed in a particular reference frame [42, 44-48, 52, 54, 55]. In particular, much attention has been paid to retinotopic or gaze-centered error patterns [42-45, 52, 54]. While we observed similar error patterns here, we found that their magnitude and directionality vary with the sensory signals available for movement planning. This variation would be difficult to explain in terms of a bias in a single neural representation. Instead, we show that a bias in the transformation of target information between retinotopic and body-centered representations can lead to gaze-dependent error patterns in both representations. This shows that the spatial pattern of reach errors need not be a good indicator of the representation in which a movement is planned. More generally, we argue that when sensory signals are used in a statistically optimal manner [57, 65-71], the same information is contained in multiple neural representations, and so there need not be a direct relationship between the behavioral output and any single representation of the movement plan.

We have proposed that a biased transformation can account for gaze-dependent reaching errors. Of course, the evidence for a purely retinotopic representation of movement planning includes the "remapping" of retinal eccentricity effects due to intervening saccades. If subjects shift their gaze subsequent to the presentation of a reach target, but before reaching, the reach error matches what would be observed in a no-saccade trial where the target is presented relative to the new gaze position rather than the original configuration [44, 72]. Since the parietal representations of retinotopic space remap with saccades [32, 73-75], these results could be explained either by a biased retinotopic representation of the movement plan or by a bias in the mapping from this representation

to the movement. Although our model does not directly address such memory delay tasks, we believe it is could be extended to address these results. It is plausible that the saccade triggers not only a remapping of retinotopic representations, but also a "re-broadcast" of those representations to other parts of the brain, and in particular to the body-centered representations in our model. During the process of remapping, the cells which code for the new location in space experience a marked increase in their firing rates, similar to that experienced when a real visual stimulus is present, albeit at a slower rate [74]. Such time-locked increases across a region of the retinotopic representation could act as the trigger for information transfer from one representation to another providing updating of errors across gaze shifts.

In this paper, we have focused on azimuthal (left-right) gaze shifts and the resulting error patterns. A vertical component to gaze-dependent errors has been observed when comparing vertically separated gaze locations [76]. These errors also follow the retinal eccentricity pattern, and are thus qualitatively consistent with our model. Gaze-dependent errors in depth were not observed in our dataset (Supplemental Fig. S4), most likely because we did not vary the gaze location in depth. When this manipulation is done, a complex pattern of gaze-dependent depth errors is observed [54], presumably reflecting the complex binocular, three-dimensional geometry of the eyes [58]. While our model might be successfully extended to account for these patterns, it would require significantly more complex representations and transformations.

Our simple model of gaze-dependent transformations ignores other possible

sources of bias. For instance, we do not address the relationship between head and eye movements. Several studies have explored the effect of head rotations on reach accuracy [77-79], and one recurring result is a bias in reaching toward the direction in which the head is oriented [77, 79]. This effect may result from a bias on the perceived midline of the head toward the center of gaze [79]. Indeed, we expect that biases in spatial transformations have a pervasive effect on movement planning. For example, the leftward bias in reach errors observed across tasks (Fig. 3) can be modeled as a rightward bias in the proprioceptive estimate of the right hand (Supplemental Section 1.2). Further, the idiosyncratic gaze-independent error patterns exhibited by individual subjects (Supplemental Fig. S2) are also likely to be explained by subject-specific biases on other sensory variables [80]. These biases might also arise due the presence of prior expectations [69], although alternate sources of bias are plausible, e.g. impoverish representations [81].

Sensory transformations can incur variance as well as bias [31, 37]. In our model, the transformation variance is due to uncertainty in the gaze direction, as proposed by Schlicht and Schrater [37]. The model thus assumes that sensory uncertainty is the only source of variability in this computation. This viewpoint is the basis of the sensory integration model of Ma et al. [82], in which computation is noise-free and neural variability serves to represent uncertainty about the external world. It is important to note, however, that in our model the degree of uncertainty in estimated gaze location was a fit parameter. This is a quantity that would be difficult to measure directly, since assessing estimated gaze location always requires comparing that estimate to other variable signals. It is therefore plausible that these terms capture both sensory uncertainty and additional

variance due to the stochastic nature of neural computation [83].

No matter the source, we argue here that the variance due to sensory transformations has an important effect on the flow of information within the reach planning circuit. In particular, we argue that these transformation "costs" make it advantageous to use multiple representations for the movement plan. First, we showed that when multiple reference frames are used, in a weighted fashion, movement variability across trials is improved (Fig. 6b). Next, we showed that our model is able to predict how the weighting of visual feedback of the hand changes with target type (Fig. 8) [31], but only when multiple representations are used. This is because the relative variability of the two representations, and hence their contribution to the output, changes with the sensory modality of the target. Note however, that the relative contributions of two signals can also depend on the degree to which they appear to come from the same source [84-86], an effect that can also account for sensory reweighting [86] and therefore could explain the portion of reweighting that is not predicted by our model (Fig. 8). Still, the decrease in planning variance with multiple reference frames and the corresponding reweighting of sensory information across trial types suggest that while a single representation might simplify the flow of information [33, 38-41], it does not make optimal use of that information for estimating the desired movement vector.

Previous studies have reported patterns of movement errors [31, 47, 49] or generalization of motor learning [87, 88] that could not be explained succinctly in a single reference frame. Such error patterns often depend on the availability of sensory signals [31,

[47, 89], as observed here. These observations have been explained in terms of task or learning-dependent changes in the underlying reference frame [86, 88] or an "intermediate" representation of movement planning [49, 87]. Here we provide a different explanation: movements are always represented in multiple reference frames, independent of the task. It is the statistical reliability of these representations which determines their relative weighting.

This model is consistent with the neurophysiological literature, where a variety of spatial representations have been observed across the reach planning network [4, 19, 20, 22]. Retinotopic coding for reaches has been observed within the intraparietal sulcus [8, 15], while other studies of parietal cortex have found head- and body-centered coding as well as "mixed" representations [10, 13, 16, 27, 90]. In the pre-motor cortex, which has strong reciprocal connections to these parietal areas [6, 91], neurons with "mixed" and hand-centered representations have been observed [18, 24, 26]. Mixed hand, shoulder, and body-centered representations have even been reported in the primary motor cortex [25]. The two separate representations of the movement plan in our model might correspond to different subsets of these cortical areas. Alternatively, the same computation could be performed using a single neural population that contains both retinotopic and body-centered components. In neurophysiological studies, such an area might appear to have a "mixed" or "intermediate" representation. Indeed, these implementations are two ends of a continuum, and the physiology seems to point to a model in which all of these cortical areas exhibit "mixed" representations, but the parietal cortex has a more retinotopic character and the frontal cortex is more hand or body-centered [19, 22, 92]. As this study has

shown, however, the ultimate answer is likely to be found not by finer assays of neural reference frames, but rather by comparing activity in these areas across tasks with different sensory information [16, 27, 33].

Materials and Methods

Experimental setup.

Subjects were seated in a simple virtual reality setup (Fig. 1). The right arm rested on the surface of a thin (6 mm) and rigid horizontal table. The left arm remained under the table. When used as a reach target, the left index finger touched the underside the table with the wrist supine. Thus, while the two index fingers could be brought into close proximity with each other, the two hands never came in contact. The location of both index fingers was monitored using an infrared tracker (Optotrak 3020, Northern Digitial, Waterloo ON). Subjects' view of their hands and arms were blocked by a mirror through which they viewed a rear-projection screen (Fig. 1a, inset). The screen and mirror were adjusted so that objects displayed on the screen with a digital projector appeared to lie in the plane of the table. The rig was enclosed in black felt and the room was darkened to minimize additional visual cues. Head movements were lightly restrained with a chin rest, and eye movements were monitored with an ISCAN Inc. (Burlington, MA) infrared eye tracker.

Task design.

Nine potential reach targets were located on the table, on a 35 cm arc centered at the point directly below the midpoint of the two eyes (Fig. 1). The targets were located at ±20º, ±15º, ±6º, ±2º, and 0º with respect to midline. Two gaze fixation points were also located on this arc at ±10º. Visual targets were displayed as 8 mm radius green disks, and fixation points were 5mm radius red disks. Visual feedback was always given in the form of an 8 mm radius disk centered on the index finger, white for the right hand, blue for the left.

All trials consisted of four steps. 1) Subjects moved their right index finger to a fixed starting location. On feedback (FB) trials, the start location was indicated with a 10 mm radius green disk, and visual feedback of right hand was illuminated. On no-feedback (NoFB) trials, neither the feedback nor the target were visible, and subjects were guided to the start location using the arrow field method, which provides no feedback of absolute hand position [31]. 2) When the right index finger came to rest within 10mm of the start location, one of the two fixation points appeared. Subjects were required to look at the fixation point and maintain fixation for the remainder of the trial. 3) When fixation was achieved, the reach target was specified. For VIS and VIS+PROP trials, the target disk appeared. In VIS+PROP trials, feedback of the left hand appeared, subjects then moved the left index finger into the target disk, and the target disk was then extinguished leaving the blue feedback disk. For PROP trials an arrow field was again used to guild the unseen left hand to the unseen target location. 4) After the target was

specified, there was a 500ms delay before an audible "go" tone was played and subjects reached to the target. On FB trials, both the feedback and the start disk were extinguished at the time of the "go" tone and remained off for the rest of the trial. Finally, subjects were required to hold the final reach position for 500 ms. Subjects practiced the various trial types before beginning the experiment.

Eight subjects (two female, six male) participated in the experiment. Subjects were right-handed, had no known neural or motor deficits, and had normal or corrected-to-normal vision. The experiment was divided into two sessions, which were performed on different days in order to minimize fatigue. One session contained only FB trials, one only NoFB trials, and session order was randomized across subjects. Each session contained six repetitions of each of the 54 conditions (3 target types x 9 targets x 2 fixation points), for a total of 324 trials (not including error trials, which were repeated). These trials were followed by a set of trials in which gaze location was unconstrained. This set consisted of six repetitions of 9 trial conditions (3 target types x 3 targets), bringing the *total* number of trials for each session to 378. The order of presentation across conditions was randomized within each repetition.

Data analysis.

For each trial, the reach endpoint was defined as the position where the movement speed first fell to 5 mm/sec. Reach targets and reach endpoints were converted into polar coordinates about an origin located directly below the midpoint of the two eyes. *Angular*

reach error is defined as the angular difference between the endpoint and target, with positive values indicting reach endpoints to the right of the target and negative values indicating reach endpoints to the left of the target. For the plots, permutation tests, and model fitting in Figures 3 and 5, the angular reach errors were corrected by subtracting off the interpolated free-gaze errors (separately for each subject and trial type) in order to minimize the effects of idiosyncratic gaze-independent error patterns while preserving the relationship between error, gaze location and target (see Supplemental Figs. S1,S2). The significance of gaze-dependent effects was tested by a paired permutation test of a main effect of gaze location [93].

Model.

Our model of reach planning describes how statistical representations of sensory inputs are used to compute a movement vector plan. Five sensory signals are potentially available, modeled as independent Gaussian probability distributions centered on the true locations, X, i.e. Gaussian likelihoods, $N(X,\sigma^2)$, with an isotropic covariance matrix $\sigma^2 I$:

$$\begin{aligned}
\text{vision of right fingertip:} \quad & p\left(f_v \mid F_R\right) \sim N(F_R,\sigma_v^2) \\
\text{proprioception of right fingertip:} \quad & p\left(f_p \mid F_B\right) \sim N(F_B,\sigma_p^2) \\
\text{vision of target:} \quad & p\left(t_v \mid T_R\right) \sim N(T_R,\sigma_v^2) \\
\text{proprioception of target:} \quad & p\left(t_p \mid T_B\right) \sim N(T_B,\sigma_p^2)
\end{aligned} \tag{1}$$

felt gaze position: $p(g \mid G) \sim N(G, \sigma_g^2)$

Here lower-case variables are sensory signals with subscripts denoting sensory modality, *v*, visual, *p*, proprioceptive. Upper-case variables are true locations with subscripts denoting reference frame, R for the retinotopic location, B for body-centered location. When a signal is not available in a given trial type (e.g., f_v in NoFB trials), the likelihood is set to the uniform distribution. The likelihood represents *variability* in a sensory signal x given the true location X. The computations in the model, however, depend on the *uncertainty* in X given x, i.e. on the posterior distributions $p(X \mid x)$. Bayes' rule relates these two distributions, as a function of X:

$$p(X \mid x) \propto p(x \mid X) p(X), \tag{2}$$

where $p(X)$ represents prior information about the location. For the most part, we assume that the prior is flat, so the posterior is proportional to the likelihood (although below we describe when the prior is not flat).

All of the computations in the model are statistical in nature, making locally optimal use of the signals required for a particular computation, assuming that those signals are mutually independent. Indeed, there are only two operations performed by the network: signal integration and addition (or subtraction). *Integration* is the process of combining the information about variable X from two signals x_1 and x_2. By Bayes rule (Equation 2) and the definition of independence, the integrated posterior is just the product of the two input distributions:

$$p(X \mid x_1, x_2) \propto p(X \mid x_1) p(X \mid x_2). \tag{3}$$

The resulting posterior is also Gaussian, with mean and variance

$$\mu_{X|x_1,x_2} = \left(\frac{\mu_{X|x_1}}{\sigma^2_{X|x_1}} + \frac{\mu_{X|x_2}}{\sigma^2_{X|x_2}} \right) \sigma^2_{X|x_1,x_2}, \qquad \sigma^2_{X|x_1,x_2} = \left(\frac{1}{\sigma^2_{X|x_1}} + \frac{1}{\sigma^2_{X|x_2}} \right)^{-1}. \quad (4)$$

In Equation 4 and below, $\mu_{X|x}$ and $\sigma^2_{X|x}$ represent the mean and variance of the distribution $p(X \mid x)$, respectively. If either of the input distributions is uniform, i.e. if a sensory signal is absent, the integrated posterior is equal to the other input. Note that the integrated variance is smaller than either of the input variances when two distributions are combined. The second operation, *addition*, is where the network computes the posterior of a variable $Z = X + Y$ from input signals x and y. In this case, the output posterior is given by the convolution of the inputs:

$$p(Z \mid y, x) = p(X \mid x) \otimes p(Y \mid y). \quad (5)$$

The result is also a Gaussian, with mean and variance

$$\mu_{Z|x,y} = \mu_{X|x} + \mu_{Y|y}, \qquad \sigma^2_{Z|x,y} = \sigma^2_{X|x} + \sigma^2_{Y|y}. \quad (6)$$

Note that for both integration and addition, the output mean is a weighted sum of the input means, with either constant (unity) weights or weights that depend on the input variances.

Given the sensory signals described in Equation 1, the model first builds internal representations of the fingertip and target locations, in both a retinotopic and a body-centered reference frame. These representations make optimal use of all available sensory signals, requiring a transformation of "non-native" signals. In computing the retinotopic representation of target, for example, the proprioceptive signal t_p must be

transformed. Since head position is fixed in our model, $T_R = T_B - G$, and the transformation follows Equation 5:

$$p\left(T_R \mid t_p, g\right) = p\left(T_B \mid t_p\right) \otimes p\left(-G \mid g\right). \quad (7)$$

Parallel transformations are used to convert f_p into a retinotopic representation and t_v and f_v into body-centered representations. This can yield two independent estimates of the same variable, which are then integrated according to Equation 3. The retinotopic representation of target, for example, has the posterior distribution:

$$p\left(T_R \mid t_v, t_p, g\right) \propto p\left(T_R \mid t_v\right) p\left(T_R \mid t_p, g\right). \quad (8)$$

Of course, in VIS or PROP trials, one of the input distributions is uniform.

The model next computes retinotopic and body-centered representation of the desired movement vector. Since the true value of the instructed movement vector is $MV = T - F$, this computation also follows Equation 5:

$$p\left(MV_U \mid t_v, t_p, f_v, f_p, g\right) = p\left(T_U \mid t_v, t_p, g\right) \otimes p\left(F_U \mid f_v, f_p, g\right), \quad U = R, B \quad (9)$$

Note, however, that both of the inputs to Equation 9 depend on gaze. If the same estimate of gaze is used to transform all sensory variables, then the inputs are not fully independent, and Equation 6 is only an approximate solution (see Supplemental Sections 2.4, 2.5 and Supplemental Fig. S7).

Lastly, the model selects a planned movement vector using one of three readout schemes. The RET and BODY readouts are just the mean values of the MV_R and MV_B

posteriors, respectively, from Equation 9. For the INTEG readout, the model integrates these two posteriors according to Equation 3, as if they were independent:

$$p\left(MV_{INTEG} \mid t_v, t_p, f_v, f_p, g\right) \propto p\left(MV_R \mid t_v, t_p, f_v, f_p, g\right)\, p\left(MV_B \mid t_v, t_p, f_v, f_p, g\right) \quad (10)$$

Note that these three readouts are the maximum *a posteriori* (MAP) estimates given the respective posteriors. In this and the previous step, the model makes use of computations that are only strictly correct when the input signals are independent (Equations 8,9), an assumption which is not always correct (see Supplementary Section 2.4)

The model described so far would produce unbiased estimates of the movement vector. The final component of the model is a bias in the transformation of target position between reference frames. We model this bias as a systematic misestimation of the gaze location due to the presence of a Bayesian prior. This prior takes the form of a Gaussian distribution, $p\left(G\right) \sim N\left(t_{prior}, \sigma^2_{prior}\right)$. The mean of the prior distribution t_{prior} is itself a Gaussian random variable with mean T_B and variance proportional to that of the target variable being transformed. The variance of the prior, σ^2_{prior}, depends on both the variance of the target distribution being transformed and a scaling factor that depends on the modality of the transformed signal (see Supplemental Table S1). Note that this prior only effects the gaze estimate used for transforming target locations; the prior is assumed to be flat in the transformation of finger locations. A more detailed discussion of the statistical properties of the transformation, and their effects on the model predictions, can be found in the Supplementary Note (Sections 2.4 and 2.5).

Since the model is composed entirely of integration (Equation 3) and addition (Equation 5) operations, all expected values in the model, including the movement vector readouts, can be written as weighted sums of the means of the initial sensory inputs, with coefficients that depend only on the variances of those signals (Equations 4,6). The trial-by-trial variances of the readouts, shown in Figure 7, can be readily computed from these coefficients and the input variances in Equation 1.

Model fitting.

The only parameters in the model are the sensory variances listed in Equation 1 and the variance of the gaze prior, σ^2_{prior}. The proprioceptive variances were set *a priori* based on previously published estimates [60]. The variance of visual signals, σ^2_v, is assumed to scale linearly with the distance of the stimulus location from the direction of gaze, and the scale factor is the first free parameter. The variance of the gaze signal, σ^2_g is the second free parameter. The variance of the gaze prior, σ^2_{prior}, is assumed to scale with variance of the target variable being transformed. Two scale factors, one for each target modality, make up the remaining free parameters. We fit these four free parameters to the average angular movement errors, after mean correction, as shown in Figure 5. The model generates Cartesian movement vectors for each trial condition, which are then converted into polar coordinates in order to obtain angular errors. The fitting procedure minimized the sum square prediction error across trial conditions using the Matlab optimization toolbox (function *fmincon*; Mathworks, MA). Optimization was

repeated 100 times with random initial parameter values, and the final parameters were largely insensitive to initial values.

Modeling other data sets.

The model predictions for previously published dataset (Figs. 7,8) relied on the model parameters fit to our own dataset with the INTEG readout. When modeling the reach errors for Buerze et al. [42] (Fig. 7), we translated their target array into the center of our workspace and rescaled the spacing of the targets in order to maintain the same azimuthal separation. The data for Figure 8 come from the gaze-fixed trials in Sober and Sabes [31], their Supplemental Figure S3.

Acknowledgements

We would like to thank Sabine M. Beurze, Stan Van Pelt, W. Pieter Medendorp, and Samuel J. Sober for generously providing their data for comparison with our model predictions.

Grants

This work was supported by the National Eye Institute (R01 EY-015679), the National Institute of Mental Health (P50 MH77970), and the McKnight Endowment Fund for Neuroscience. L.M.M.M. was supported by a graduate fellowship from the National Science Foundation.

Supplemental Information for Chapter 1: Sensory transformations and the use of multiple reference frames for reach planning.

Figures, equations, and tables in this Supplementary Note are numbered with a preceding "S", e.g. **Figure S1**. Any references to figures without the leading "S" refer to a figure in the main text.

1. Additional Data and Analyses

1.1. Cartesian reach errors for an example subject

The pattern of reach errors for each target, gaze location, and trial type (Fig. 1) are shown for a sample subject in Figure S1 (same subject as Fig. 2, main text). Here we highlight several key features of this subject that are typical of what we observed across subjects and that can also be observed in the mean angular errors in Figure 3 of the main text. First, while the error patterns differed across trial types, they vary smoothly as a function of target location and show a systematic dependence on gaze location. During trials with a visual target (VIS and VIS+PROP trials) the reach endpoint is relatively biased away from the point of fixation. This effect is most pronounced in reaches to

targets that lie between the two gaze locations, suggesting a sub-linear relationship between reach bias and the retinotopic distance to the target (c.f. the retinotopically aligned insets in Fig. 3, main text). The magnitude of this gaze-dependent effect depends on the availability of sensory signals. Specifically the effect is smaller when either visual feedback of the reaching hand or a proprioceptive target is available. In contrast, reaches to PROP targets are relatively biased toward the fixation point, although the effect for this sample subject is small and variable. These gaze-dependent effects show the same directional trends across all subjects and reached significance on an individual subject level in most tasks for all eight subjects. In addition to these gaze-dependent effects, this subject exhibits gaze-independent biases (i.e. common effects for both gaze locations) that vary across targets and trial types. In order to measure these biases, we had subjects make additional reaches to three of the target locations in a gaze-free condition (green lines in Fig. S1). These gaze-independent effects are much more idiosyncratic across subjects, as discussed in the next section.

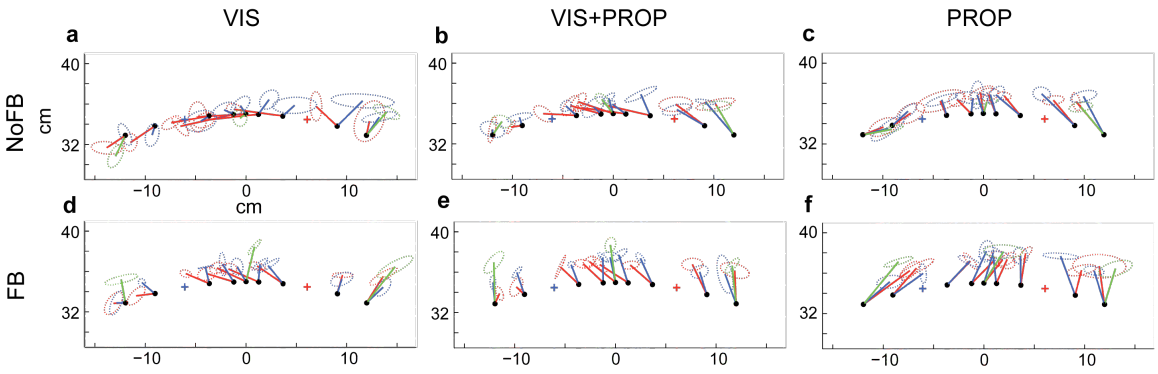


Figure S1: Mean reach errors across the workspace for an example subject. Lines show mean reach error for an individual target: red is fixation right, blue is fixation left, and green is gaze-free reaches. Ellipses show standard error of the mean.

1.2. Gaze-independent effects and inter-subject variability in reach errors

Figure S2 shows the pattern of mean errors for each subject during the gaze-free reach condition, in which no fixation point was presented. While there is substantial inter-subject variability in these error patterns, for individual subjects they vary smoothly across the workspace and are similar across trial types. Consistent with previous reports [46, 47, 94], many of these error patterns resemble a constant error vector in a polar coordinate frame centered about a point near the starting location, head, or shoulder. While such error patterns could be well-described in terms of spatial biases, the idiosyncratic nature of these errors across subjects would make it difficult to incorporate them into the model. We chose instead to focus the gaze-dependent errors, which followed more consistent patterns across subjects.

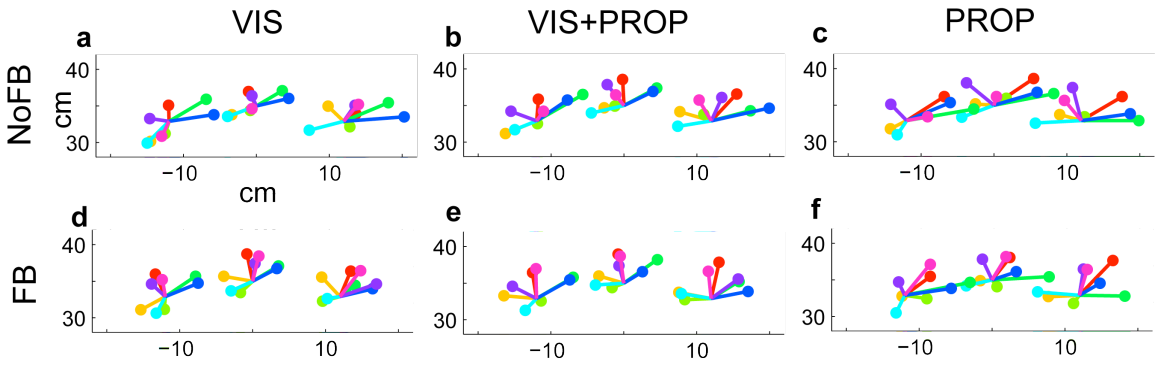


Figure S2: Mean gaze-free reach error for all subjects. Subjects made gaze-free reaches to targets located at -20, 0, and +20 degrees from the midline. Vectors show average error for each target. Each color shows reach errors for a different subject, colors are consistent across tasks.

In order to isolate the gaze-dependent errors (Fig. 3, main text), we subtracted a linearly interpolated estimate of the mean gaze-free angular error from the angular errors in the gaze-fixed trials for each subject. Figure S3 shows the mean (across subjects) gaze-

free errors and uncorrected gaze-fixed errors. A comparison of the mean corrected (Fig. 3) and uncorrected (Fig. S3) errors shows that the key trends in the data were present before correction. The gaze-free reach errors accounted for a large portion of the inter-subject variability in gaze-fixed errors, as can be seen by comparing the error bars across these two figures.

One difference between the corrected and uncorrected errors is that when the gaze-free reach errors are subtracted off (corrected), a leftward bias in reach error is present across all trial types (mean 0.69 ± 1.58 deg., present in all trial types with no significant difference between trial types; see Fig. 3). A potential explanation for these errors is the presence of a bias on the internal estimate of the right hand in gaze-fixed conditions. Indeed, we can largely capture this effect in our model by including a bias to the proprioceptive estimate of the right hand. When fit to the data in Figure 3, we obtain a best-fit value of a 0.63 cm rightward bias in proprioception with a resulting mean angular error of 0.80 degrees across all trial conditions. The effect of this bias is largely orthogonal to the gaze-dependent effects, as evidenced by the fact that fitting with the bias parameter leads to only slight changes in the other best-fit parameter values. Of course the effects of this bias is only observed when gaze is constrained, suggesting either that the bias is itself a result of the gaze constraint or that when the gaze is free the effect is absent or highly attenuated. This analysis exemplifies how the model can be expanded to incorporated additional error effects, allowing the development of a more comprehensive model of movement planning.

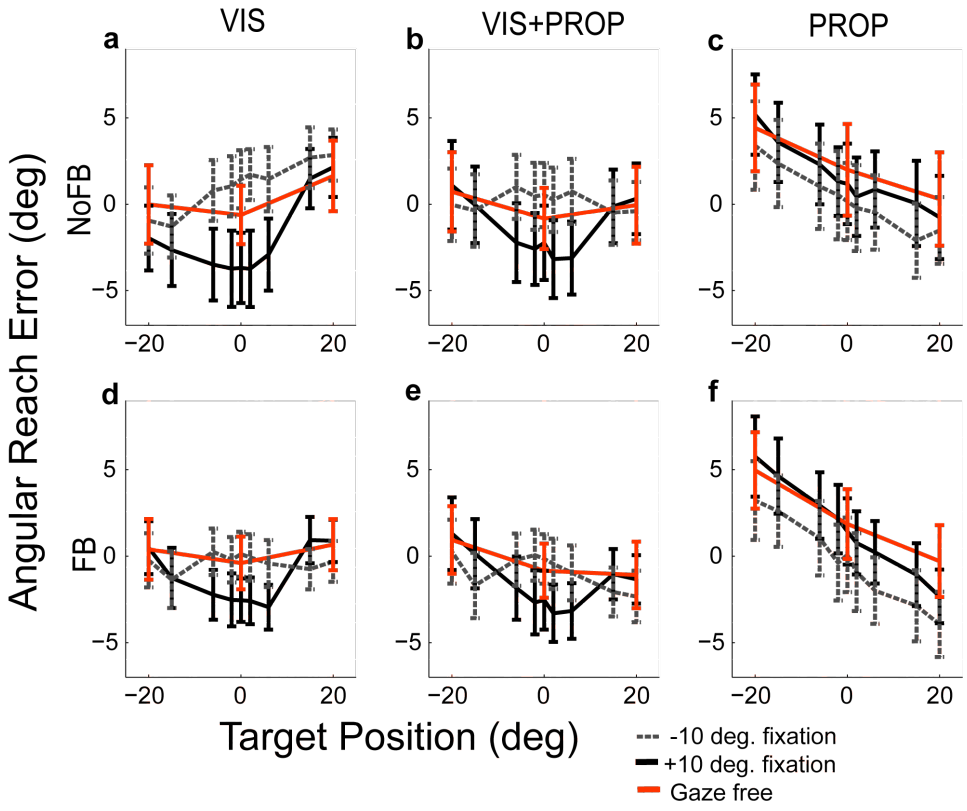


Figure S3: Average angular reach error across subjects without gaze-free biases subtracted off. Error bars show standard error of the mean.

1.3. Errors in depth

This paper focuses only on reach errors in the azimuthal angle that result from gaze displacements from the target location along the azimuthal axis (targets and gaze locations were all equidistant from the midpoint of the eyes). A previous study of reaching to visual targets found error patterns similar to those observed here for gaze displacements in both azimuth *and* elevation [76], suggesting that our model would generalize to account for angular effects along any axis. Other researchers have observed gaze-dependent errors when the fixation point is displaced in *depth* relative to a visual target [54]. However, these error patterns are relatively complex and are not likely to be explained in terms of a simple misestimation of gaze angle. We suspect that this

difference is due in large part to the greater complexity involved in interpreting vergence and other depth cues, compared to angular cues [58].

It is possible that by focusing on angular errors, we could be missing significant gaze-dependent effects in the depth components of the reach errors. A plot of the radial component of the reach errors for all trial conditions (Fig. S4) shows no significant difference between trials with left and right gaze fixation (solid and dashed lines). This observation is supported by a 4-way ANOVA of the radial errors for the gaze-fixed trials (Target Location x Gaze Location x Target Type x Feedback). The only significant effects are the main effects of Target Location ($p=0.005$), seen in the sloping lines in Figure S4, and in Target Type ($p<0.0001$). In particular, there is no effect of gaze location on the radial reach errors. In contrast, the same ANOVA performed on the angular errors in Figure S3 indentified interactions between Target Type and both Gaze Location ($p<0.001$) and Target ($p<0.001$), as well as the main effects seen in radial errors (Target Location, $p=0.03$; Target Type, $p<0.0005$). These results show that our focus on azimuthal errors allowed us to study angular gaze-dependent errors fully while isolating them from the more complex errors that arise from gaze displacement in depth. Further, our model is consistent with the data, both in its ability to capture the experimental data (Figs. 5-6, main text) and in its prediction of no radial gaze-dependent errors due to angular gaze-displacement from the target in our experimental design.

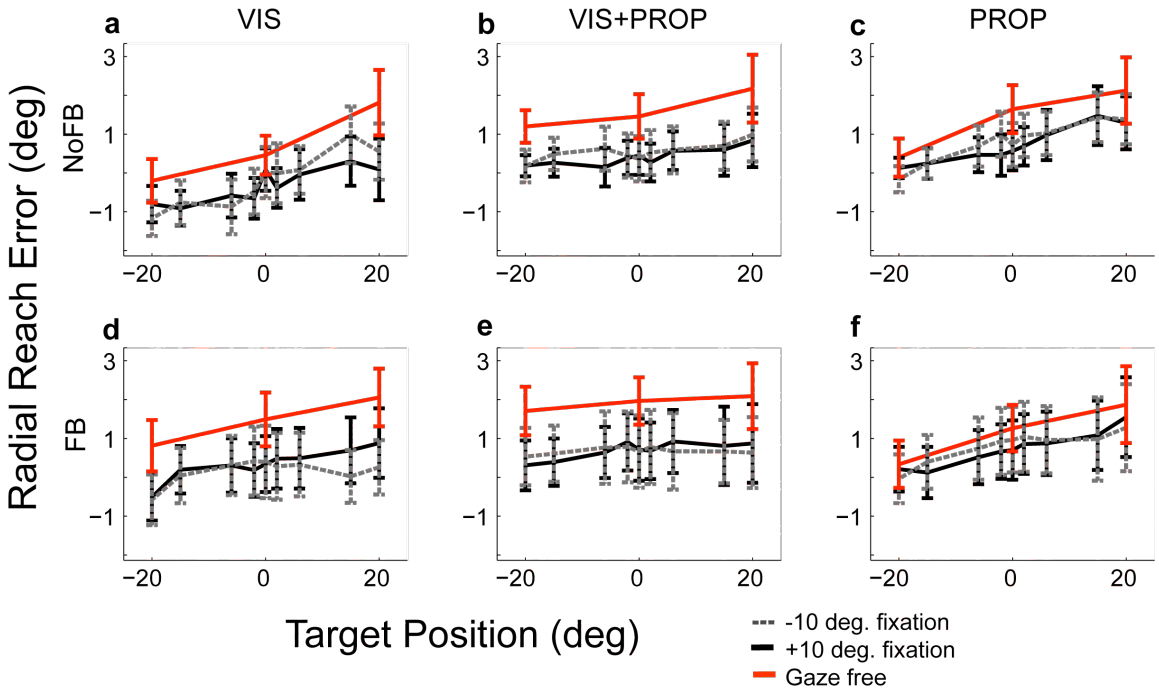


Figure S4: Average radial reach error across subjects for the two gaze conditions. Error bars show standard error of the mean.

2. Additional Model Details and Analyses

2.1. Transformation variance and the relative variability of representations

Here we show how the cost of a sensory transformation between retinotopic and body-centered representations (Fig. 4, main text) affects the relative uncertainty of integrated position estimates in these two representations. We begin with three assumptions: 1) both visual and proprioceptive information about the position are available and they are independent given the true location, 2) both signals are transformed into the other ("non-native") reference frames, and 3) the two signals are optimally combined [65] (minimum variance) in each reference frame. If σ_v^2 and σ_p^2 are the

variances of the visual and proprioceptive signals, and if σ^2_{xform} is the variance injected by the transformation (c.f., Equation 6, main text), then the resulting variances of the two combined position estimates (retinotopic and body-centered) are

$$\sigma^2_R = \left[1/\sigma^2_v + 1/\left(\sigma^2_p + \sigma^2_{xform}\right)\right]^{-1} = \left(\sigma^2_v\sigma^2_{xform} + \sigma^2_v\sigma^2_p\right)/\left(\sigma^2_v + \sigma^2_p + \sigma^2_{xform}\right) \quad \text{(S1)}$$

$$\sigma^2_B = \left[1/\sigma^2_p + 1/\left(\sigma^2_v + \sigma^2_{xform}\right)\right]^{-1} = \left(\sigma^2_p\sigma^2_{xform} + \sigma^2_v\sigma^2_p\right)/\left(\sigma^2_v + \sigma^2_p + \sigma^2_{xform}\right) \quad \text{(S2)}$$

From the expressions on the right, it is evident that the least variable representation is the one with the smallest variation in its associated sensory input. For example, when visual variability is smaller than proprioceptive variability, the retinotopic representation will be have lower variance than the body-centered representation. When one of the sensory inputs is missing, then the native reference frame of the sensory input which is present will always have lower variance. For example, if there is no visual input, then $\sigma^2_B = \sigma^2_p$ while $\sigma^2_R = \sigma^2_p + \sigma^2_{xform}$. Thus, the optimal representation of target and hand position depends both on the relative reliability of the two sensory modalities and on which sensory inputs are available.

2.2. Fit values of sensory and motor variances

The parameters of our model are terms describing sensory variability and internal uncertainty (Table S1). These parameters are either variance values (e.g. σ^2_p) or

parameters from which variances are computed (e.g. the slope of σ_v^2 with retinal eccentricity). Interpreting the latter can be less intuitive, and so Figure S5 provides a direct visualization of the key variances for each model fit. In all cases, these variances are scalar; the model assumes isotropic variability of all sensory signals, a simplification [60, 70, 95].

	RET Readout	BODY Readout	INTEG Readout
Vision $\sigma^2_v = m \mid T_b - G \mid + b$			
m (cm^2/cm)	0.001*	0.086	0.029
b (cm^2)	0.001	0.001	0.001
Proprioception			
Right hand (cm^2)	0.30	0.30	0.30
Left hand (cm^2)	0.90	0.90	0.90
Gaze (cm^2)	1.16	0.002	1.13
Prior $s \cdot \sigma^2_{tx}$			
Ret. -> Body s	3.05	0.023	12.8
Body -> Ret. s	20.0*	3.13	4.54

Table S1: Variances for RET, BODY and INTEG readout models. Fit variances are shown in black. Variances that were set before fitting are shown in grey. * Value went to limit of fit range.

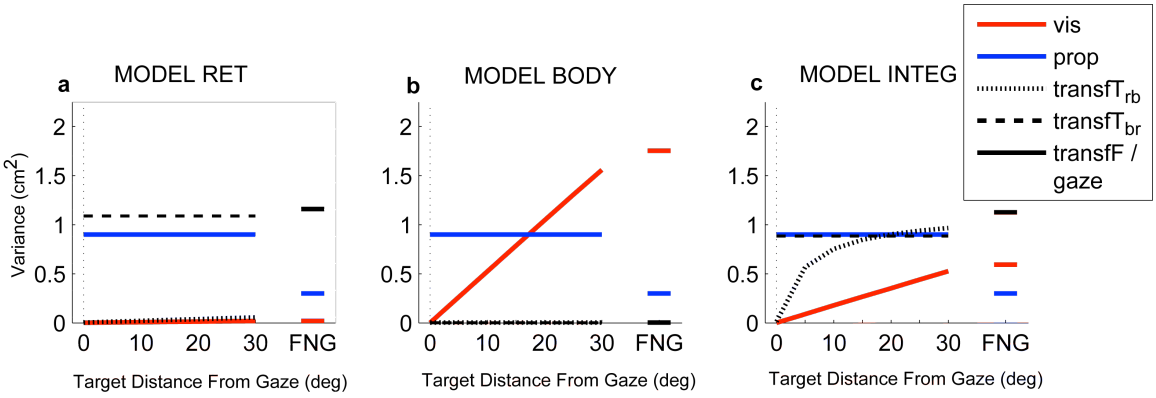


Figure S5: Best fit variance of sensory inputs for each output model. Variance of visual and proprioceptive signals for target location and transformation variances for target information are plotted as a function of the distance between target and the center of gaze. Since only one start location was used the variances of the initial hand signals are constant.

Proprioceptive Variance: The proprioceptive variances were determined *a priori*. Values were roughly estimated from the values measured by van Beers et al.[60], who deconvolved the behavioral variability measured in inter-sensory alignment tasks at three different locations in the workspace in order to estimate the underlying sensory variances.

Visual Variance: The slope of visual variance σ_v^2 on the retinotopic distance of the stimulus was the first fit parameter. For this parameter, the INTEG fits appear most consistent with the existing literature. van Beers et al.[60] inferred the variability of the visual sense of hand location in same study discussed above. They found azimuthal standard deviations in the range of 0.20–0.46°, which translates to a Cartesian variance σ_v^2 between 0.02–0.10 cm^2 in our workspace. Since gaze location was not controlled in these studies, direct comparison to our fits is difficult. Nonetheless, this variance range matches those we found for the INTEG parameters (Fig. S5c) when the targets are located approximately 0.9°–5.0° from fixation, a plausible range for the gaze-free alignment trials performed in the van Beers studies. Schlicht and Schrater [37], in building a model very similar to our own, inferred values of σ_v^2 from previous two-point discrimination studies. For the range of targets used in our study, their parameters would yield σ_v^2 in the range of 0.003-0.12 cm^2, a range 4-5 times smaller than we what we inferred with the INTEG fit. This difference may well be due to differences inherent in a two-point discrimination task versus a more difficult localization task. In any case, the INTEG fits compare much more favorably than those for the other readout models: the

values from the BODY fit were significantly larger (Fig. S5b), and the values from the RET readout were implausibly small (Fig. S5a).

Transformation variance: The total transformation variance, i.e. the uncertainty added to a sensory signal due to the transformation, shown in the black lines of Figure S5. It is somewhat difficult to assess these values since there have been no previous attempts to measure these quantities. Nonetheless, Schlicht and Schrater [37], inferred this quantity (which they called "coordinate transformation uncertainty", or CTU) quite indirectly from a study of saccade variability [96]. When converted into our workspace, their values for the transformation variance ranged from 0.024-0.65 cm^2. This is comparable to, but slightly lower than, the values we obtained with the INTEG fit, which ranged from 0.6 to 1.0 cm^2 over the targets used in our study.

It is important to note that for the RET and BODY fits, the transformation variance is not a well-constrained variable. To see why, first consider the fact the transformation bias (i.e. the effect of the prior expectation of gaze position) depends only on the *relative magnitudes* of the gaze and prior variances, σ_g^2 vs. σ_{prior}^2, not on their absolute magnitudes (c.f. Equation 4). Second, the uncertainty of a transformed signal is the sum of the original sensory variance and the total variance of the transformation (c.f. Equation 6). Since only one representation is used in the RET and BODY readouts, only this sum matters, not the individual terms. Thus, for example, in the RET readout panel of Figure S5, it is the sum of the blue and black lines that matter, not the absolute values

of the two lines. Finally, while the transformation variance in the model is formally equivalent to the uncertainty in gaze direction, we estimated its value simply by fitting to the reach errors. It is therefore plausible that this term captures both sensory uncertainty and additional variability that arises during the transformation, e.g. due to the stochastic nature of neural computation [83].

2.3. Reference frame weighting in the integration (INTEG) readout

In the main text, we argue that the fits obtained with the INTEG readout are evidence for the presence of multiple reference frames for reach planning. In making this argument, it is important to consider the relative contributions of the retinotopic and body-centered movement vectors to this readout. In particular, we want to show that both representations contribute to the movement plan across trial conditions, as opposed to a single-representation plan with a task-dependent choice of reference frames. To address this issue, we plot the relative contribution, or "weighting", of the retinotopic reference frame for each trial type, as a function of retinotopic target distance (Fig. S6). The value of the weighting can vary between 0 (BODY readout) and 1 (RET readout). For all trial conditions, the weighting is in the range of approximately 0.2-0.7. Thus, the INTEG readout does make significant use of both reference frames on any given trial.

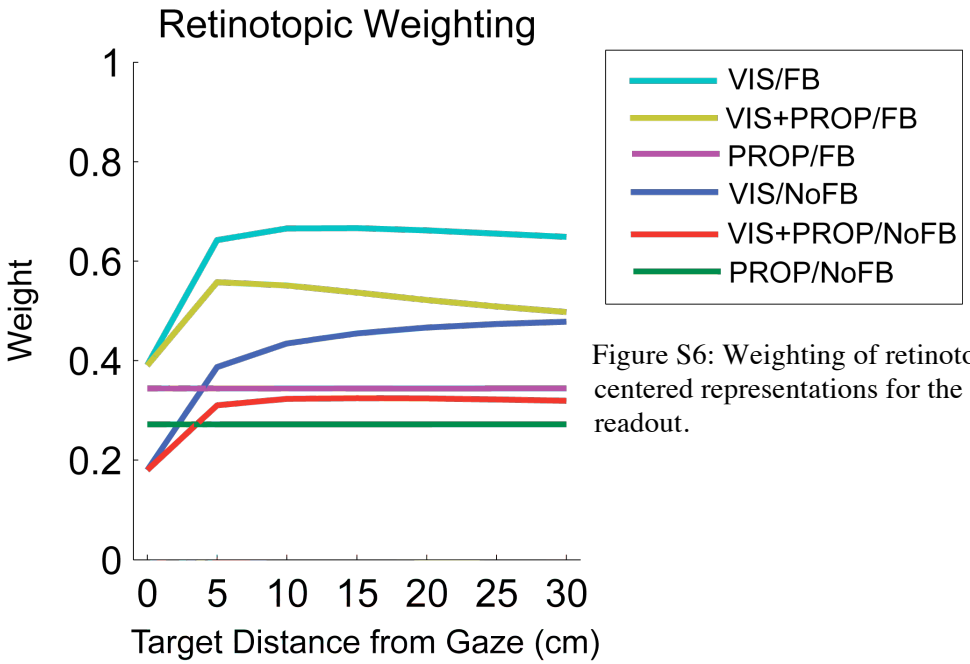


Figure S6: Weighting of retinotopic vs. body-centered representations for the INTEG model readout.

2.4. Variability, uncertainty and the assumption of independent inputs

In the *Methods* section of the main text, we distinguish between the trial-by-trial *variability* of a sensory signal x given the true value X and the model *uncertainty* of X given the signal x. For the initial sensory inputs the former is given by the variance of the likelihood (Equation 1), the latter by the variance of the model posterior, and they are simply related by Bayes' rule (Equation 2). When the prior is assumed to be flat, these two variances are equal. However at later stages, the model uncertainty about X given x, i.e. the variance of the posterior, will not necessarily equal the trial-by-trial variance of the mean of the posterior, $E(X \mid x)$. There are two reasons for this: the presence of priors in the model and the assumption of independence between local input signals. We

will discuss the affect that priors have on the trial-by-trial variability in the next section. Here we briefly discuss the assumption of independence.

The core computational elements in our model, Equations 3 and 5, implicitly rely on the assumption that the input signals are independent, despite the fact that they sometimes are not (e.g. Equations 8, 9). The assumption of independence was motivated in part by recent models of optimal sensory integration (Equation 3) and transformation (Equation 5) in neuronal populations [82, 97]. In these models, the uncertainty of each input population is encoded independently in the gain and/or neural variability of the population, and trial-by-trial correlations between these inputs are ignored. This is a simple and attractive model, as higher-dimensional correlations need not be represented and each brain area need only consider the signals it is currently receiving in order to perform its computation. In principle, however, a downstream area could learn such correlations and re-weight the inputs accordingly. Therefore we implemented a model that took all correlations into account (i.e. a model that performed true MAP estimation). While this model was roughly able to capture the mean errors observed in our data, the changes in variance across tasks did match our observed variances (data not shown). This result provides support, albeit quite indirect, that the brain performs local inference using approximations of independence.

2.5. The statistics of the sensory transformation and the optimality of multiple reference frames

A principal result of this paper is that the use of multiple reference frames for motor planning leads to a reduction in the variability of the final movement plan, as shown in Figure 6b in the main text. Here we show that, while the relative variability of the three readout schemes is highly dependent on the statistical details of the sensory transformation, for all options that we considered the INTEG readout remains the best overall performer (Fig. S7). In particular, we consider how the degree of correlation between various transformation-related signals affects model output variability. As noted in Section 2.4 above, individual computations in the model assume that local inputs are independent, and so estimates of uncertainty within the model are unaffected by correlations between sensory signals. Since the weighting of sensory signals depends only on these estimates of uncertainty (Equation 4), the MAP estimates of location, and thus the model fits (Fig. 5), are also unaffected by the presence or absence of correlations between sensory signals. Of course the actual trial-by-trial variability of the output (Fig. 6) does depend on these correlations. We consider two specific cases here.

Gaze signals: The transformations between retinotopic and body-centered representations require the addition or subtraction of the gaze location. Four such transformations are performed: the retinotopic to body-centered transformation (R→B) and the body-centered to retinotopic transformation (B→R), for both the target and initial finger locations. Since these transformations are likely to take place within different

neural circuits, they are also likely to be at least partly uncorrelated from each other across trials. For the model, four scenarios were considered: *4-estimates*, in which each transformation has a separate independent and identically distributed (i.i.d.) gaze signals (as per Equation 1, Fig. S7a); *2-estimates*, in which the R→B and B→R transformations rely on separate i.i.d. gaze signals (Fig. S7b); and *1-estimate*, in which all four transformations use the same gaze signal (Fig. S7c).

Target signals: The gaze-dependent reach errors in the model are due to the presence of a Bayesian prior on gaze location. This prior is only applied during the transformation of the target location. While this choice was made in based on our behavioral data (Fig. 3), it can be motivated using the same "local statistics" argument presented in the previous section. Specifically, we hypothesize that the prior arises within the neural circuits performing this transformation in response to the long-term statistics of gaze and target locations. The prior takes the form $p(G) \sim N(t_{prior}, \sigma^2_{prior})$, where t_{prior} is itself a Gaussian random variable. Consistent with the local statistics argument, t_{prior} is drawn from the same distribution as t_v for R→B transformations and the same distribution as t_p for B→R transformations. There are two ways to implement this model: *independent prior*, where t_{prior} is independent from t_v and t_p (Fig. S7, solid lines); and *dependent prior*, where the same variable is used for t_{prior} and t_v or t_p (Fig. S7, dashed lines). Note that in both cases the trial-by-trial variability of the mean of $p(G)$ is determined by the target variance, σ_v^2 or σ_p^2. In contrast, the uncertainty in the prior

σ^2_{prior} is proportional to the target variance (gain factors listed in Table S1). This is another example in which variability and uncertainty diverge in the model.

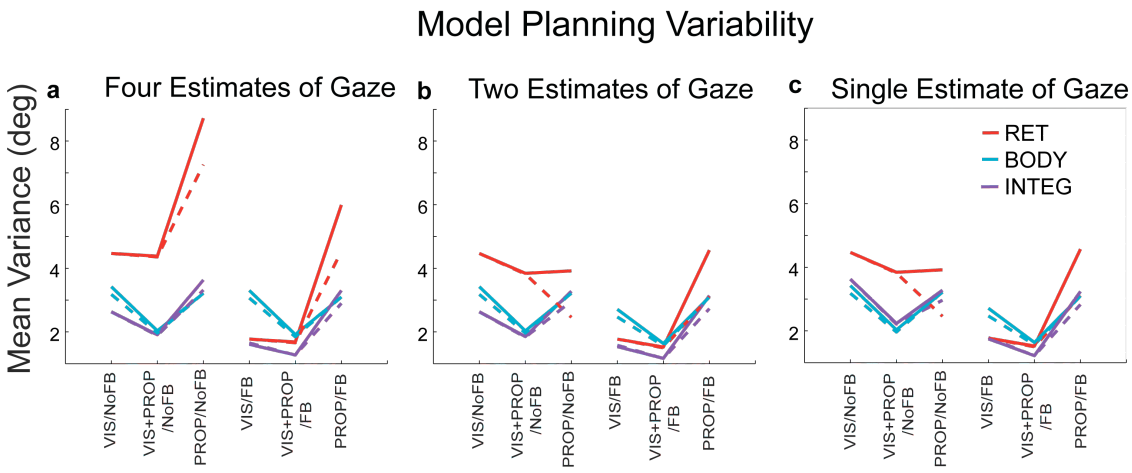


Figure S7: Predicted planning variability across tasks for models with different degrees of correlation between the estimate of gaze used to transform between reference frames. Solid lines show model predictions when prior on gaze is independent of target information, dashed lines show model predictions when the prior is the estimate of target location. a) Shows model predictions when the same estimate of gaze is used to transform all sensory signals. b) Shows model predictions when the estimates of gaze used to transform information from retinotopic to body-centered representations or from body-centered to retinotopic representations are independent. c) Shows model predictions when the estimates of gaze used to transform each sensory input (visual or proprioceptive target or hand position) are all independent.

Figure S7 shows the predicted output variability for each combination of scenarios just described. For every scenario, the integrated movement plan (purple curves) performs markedly better than either of the unimodal readouts when visual feedback of the reaching hand is available (FB conditions). When no feedback is available, the body-centered readout (blue curves) outperforms the integrated readout for scenarios with a single gaze estimate (Fig. S7c). Note however that the difference between the variances is quite small in these conditions. More generally, we expect that the true answer would lie somewhere in between these extremes, i.e. that different representations of the same external stimuli would be correlated, but not identical. In this

scenario, the integrated and body-centered readouts will perform comparably for no-feedback reaches, while the integrated readout will perform best for reaches with feedback. Notably, the retinotopic readout is generally a good deal more variable than the others.

3. Experimental Investigations into the Origin of the Gaze Bias

A key element of our model is a bias in the internal estimate of gaze location used to transform target information. While we show in the main text that this bias allows us to predict complex patterns of reach errors, we did not provide an explanation of how this bias might arise. Here we describe two experiments that tested possible explanations for a gaze bias and another experiment that quantitatively measured the bias independent of the approach used in the rest of the paper.

3.1. Saccadic planning and gaze bias

We first tested the hypothesis that the bias in estimated gaze location results from the (partial) updating of the gaze estimate due to a "covert" saccade planned toward the pending reach target. Neurons in the posterior parietal cortex have been shown to remap their patterns of activity in order to account for a shift in retinotopically represented spatial variables before a pending saccade [74]. If subjects plan (but do not execute) saccades to the reach target location, a partial shift in representation could occur. This

shift would mimic a bias in estimated gaze direction. If this were the case, then when subjects make a saccade to an alternate location, it should change the direction of the effective gaze bias, thereby disrupting the observed pattern of gaze-dependent reach errors.

To test this idea, we designed an experiment in which subjects were required to saccade to a cued location at the onset of the reach. Four subjects participated in this experiment. The saccade target could either be the same as the reach target or a position on the opposite side of the initial gaze location from the target (Fig. S8a-c). Trials were a modified version of the VIS/NoFB trials from the main text. Two reach targets were used, 10 degrees to the right or left of midline. For each reach target, two initial gaze locations were used, 10 degrees to the right or left of the target. For each reach-target/gaze combination, subjects were instructed to make a saccade to a cue location 10 degrees from the fixation point. The cue was either at the reach target or 10 degrees from the fixation point in the opposite direction (Fig. S8b,c). This resulted in a total of eight different trial conditions, and subjects completed 20 repetitions of each. To ensure that the saccade coincided with the start of the reach, subjects were required to maintain fixation until they started moving and to fixate the new eye target before their fingertip moved 40% of the nominal reach distance. Subjects were required to make single continuous movements to the reach target in order to successfully complete a trial. This experimental design allowed us to measure the effect of making a saccade to or away from the reach target across multiple target and gaze locations.

If this saccade hypothesis were correct, we would expect to see a reduction or reversal in the magnitude of gaze dependent errors when subjects saccade to the cue opposite the reach target and no change in gaze dependent errors when subjects saccade to the reach target. However, no such difference was observed (Fig. S8d). Indeed, for both cue locations, trials with initial fixation to the left of the reach target (filled markers) had rightward errors, and trials with initial fixation to the right of the reach target (open markers) had leftward errors. This is the retinal eccentricity effect observed in the VIS/NoFB condition in the main paper (Fig. 3a), and it is not reduced by peri-reach saccade away from the target as predicted by this hypothesis. These results suggest that saccade-related remapping of retinotopic representations does not lead to a bias in the internal estimate of gaze location. However, it is also possible that we were unable to sufficiently control the time of the saccade. For example, subjects could have been covertly planning a saccade to the reach target during the critical reach planning period, even though they were later able to execute a saccade in the opposite direction. The latter explanation is not entirely infeasible: since subjects found it very difficult to initiate a saccade to one location while reaching to another, somewhat loose timing criteria for the saccade were required in order to obtain reasonable success rates.

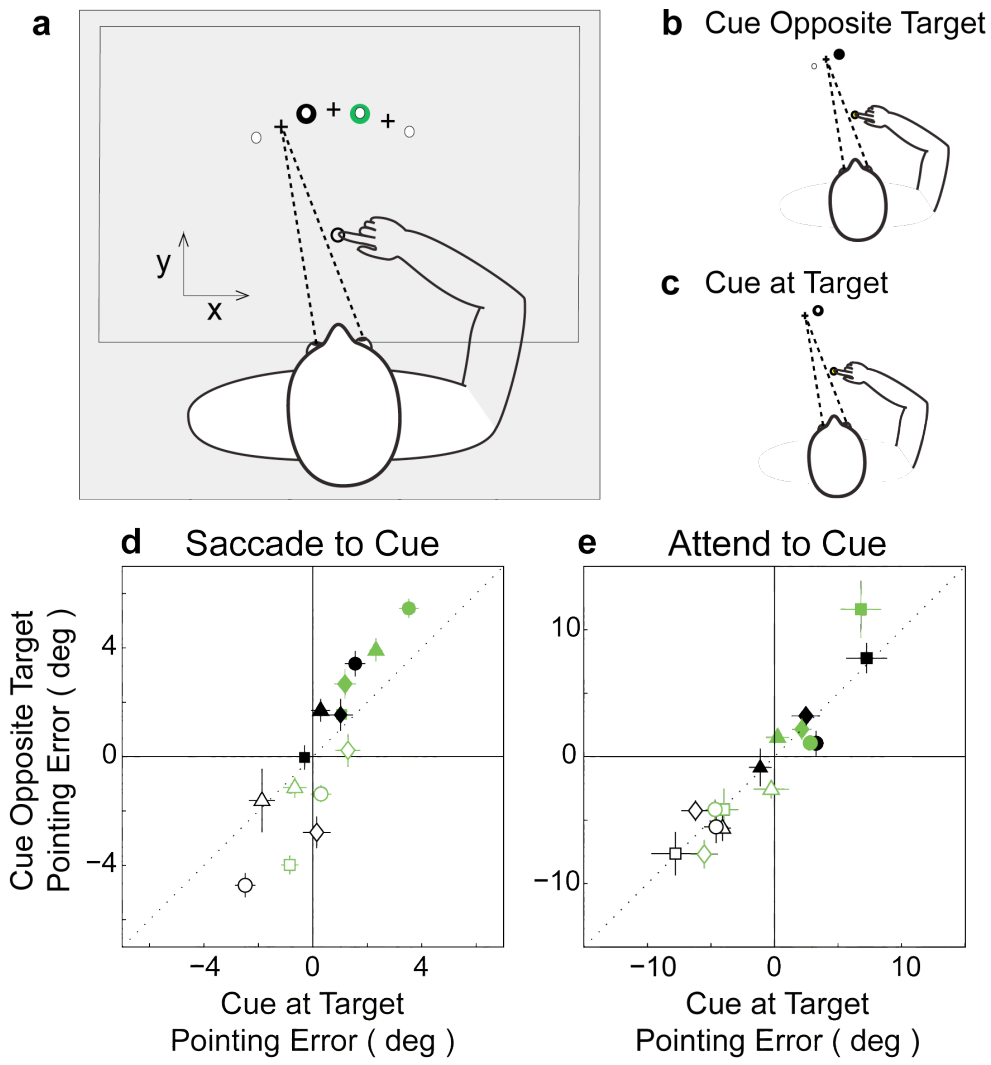


Figure S8: Effects of saccade and attention on gaze bias. a) Target configuration for both experiments: • left target, • right target, + fixation points, o saccade target and cue positions. For each reach target, the two nearest gaze locations were used. For each gaze location and target, the cue was located opposite the target (b) or at the target (c). d,e) Magnitude of pointing errors for cue opposite target versus cue at target. (d) Saccade task, (e) Attentional task. Colors represent target positions in (a). Filled symbols indicate fixation to right of the target, hollow symbols indicate fixation to left of the target. Each marker shape corresponds to data from one subject. Error bars show standard error of the mean.

3.2. Reach planning and the locus of attention

We next tested a similar hypothesis, that having the locus of attention at the reach target is the source of the gaze bias. It has been shown that the presence of a salient

peripheral visual cue can bias the report of target located away from that cue [98, 99]. This "attentional repulsion effect" resembles the retinal-eccentricity effect, and is consistent with a shift in the internal estimate of gaze toward the attentional locus. We tested whether manipulating attention could alter gaze-dependant reach errors using an experimental paradigm similar to that described in the previous section. Four subjects participated in this experiment. The target, gaze location, and cue configurations used in this experiment were identical to those described above. However, rather than having subjects saccade to a cued location, subjects were required to attend to the cue in order to detect reach "go" and "abort" signals. In this experiment, no go tone was used. Instead, the white attentional cue turned a pale green (from [1 1 1] to [.5 1 .5] in RGB space) to indicate "go". Subjects then had 800 ms to start the movement or the trial was considered an error. In order to ensure that subjects continued to pay attention to the cue after reach initiation, an "abort" cue was issued during the movement on 35% of the trials (the attentional cue turned pale red, [1 .6 .6] in RGB space), at which point subjects were required to move their hand back towards the start position within 800 ms or the trial was considered an error. Subjects quickly learned to minimize error trials. As in the previous experiment, the attentional cue was located either in at the reach target (Fig. S8c) or on the opposite side of the gaze location from the reach target (Fig. S8b). Subjects completed 20 repetitions of all target, gaze and cue combinations.

If attentional locus biases the estimated gaze location, then we would expect that the direction of gaze-dependent errors would reverse when subjects attended to the cue opposite the reach target. Instead, we observed the same retinal eccentricity pattern in

with both cue locations (Fig. S8d). While subjects displayed large gaze-dependent error patterns consistent with those observed above and in the main text, the error patterns for the two attention conditions were almost exactly matched. We conclude that either attention to the reach target is not the origin of the gaze bias or that this paradigm was unable to sufficiently divert attention from the reach target.

3.3. Measuring gaze bias

The two experiments above seem to exclude a clear role for the saccadic or attentional systems in the origin of a gaze bias. A third possibility is that the bias estimated gaze direction arises from a prior expectation that the reach targets tend to be foveated. This expectation could reflect the fact that eye and hand movements tend to be tightly linked [59, 61, 62, 100]. If this expectation acted as a Bayesian prior [63], it would bias the internal estimate of gaze direction used to transform target information toward the target. It is not yet apparent to us how the existence of such a prior could be convincingly proved or disproved. Instead, we performed an experiment designed to detect and quantify any reach-related gaze bias, using measures different from those in the main paper.

Subjects viewed a transient visual stimulus and were asked to report whether the stimulus was located to the right or left of “straight ahead”. In order to make this assessment, subjects must use the internal estimate of gaze direction to compare the

retinal stimulus to the head- or body-based sense of straight-ahead. Any bias in the internal estimate of gaze should translate to a bias in the estimated location of the stimulus with respect to straight-ahead (Fig. S9b,[64]), providing an indirect measure of gaze bias. We performed this perceptual task in the context of reaching trials, with the expectation that the internal estimate of gaze location would be biased toward the reach target. However, it is known that a peripheral distracter can bias the reported location of a flashed cue, even when the distracter is not a movement target [98, 101]. Therefore, we also tested whether a target-like cue might be enough to bias reporting of straight ahead, even when no reach was required. This experiment was thus performed with both reach and no-reach trials. During reach trials, subjects both reached to a peripheral visual target and reported the location of a bar that was flashed when the hand was ~40-60% of the way to the reach target. In no-reach trials, subjects were presented with the same scene as in the reach trials, but their hand remained in the start position throughout the trial. Trials were randomly interleaved, but subjects were informed before each trial whether it would be a reach or no-reach trial, and the target color was different for the two trial types. For each of these trial types, 22 bar locations, two gaze locations, and four targets were used, for a total of 176 trials. For each gaze location and reach target, we fit a psychometric curve to the categorical response data as a function of bar location (left or right). From the inflection point in this fit, we obtained an estimate of localization bias. Since eye position has been shown to effect subjects' report of straight ahead [79], the two gaze locations were located symmetrically 3 deg on either side of midline (Fig. S9a), and the bias estimates from the two gaze locations were averaged. We thus obtained an estimate of localization bias for each subject, trial type, and target. In order to remove

idiosyncratic biases in perception of straight ahead, we removed the grand mean for each subject before averaging across the six subjects who participated in this experiment. Two of six subjects showed such extreme leftward biases in the estimate of straight ahead that insufficient "rightward" responses were given in order to obtain a good psychometric fit. For these subjects trial conditions with insufficient data were excluded from the group average. This resulted in the exclusion of all data collected with one of the gaze locations for the first subject, and the reach data for one gaze location for the second subject. Because we subtracted off the subjects' grand mean before averaging across subjects, the bias caused by excluding one gaze position did not affect the measure of the bias due to target position.

We found that the presence of peripheral visual targets induced a bias in the localization of visual stimulus with respect to straight-ahead (Fig. S9c). For both reach and no-reach trials, this bias is consistent with a bias in estimated gaze location toward the target. However, the effect magnitude is roughly half of what would be needed to fully account for the gaze-dependent reach errors we observed in the main paper (compare Fig. S9c to Fig. 3d). This difference is likely due to the differences in the two measures used. Specifically, the gaze bias used in the model incorporates both the position of the eyes and the position of the head on the body, while the perceptual effect measured here only took into account the position of the eyes (subjects were asked to report whether they perceived the line to be to the right or left of a line coming directly out of their nose). Finally, the fact that this effect is independent of the intent to reach would appear to suggest either that subjects were covertly planning reaches even during

no-reach trials or, alternatively, that some sort of attentional mechanism is responsible for the effect.

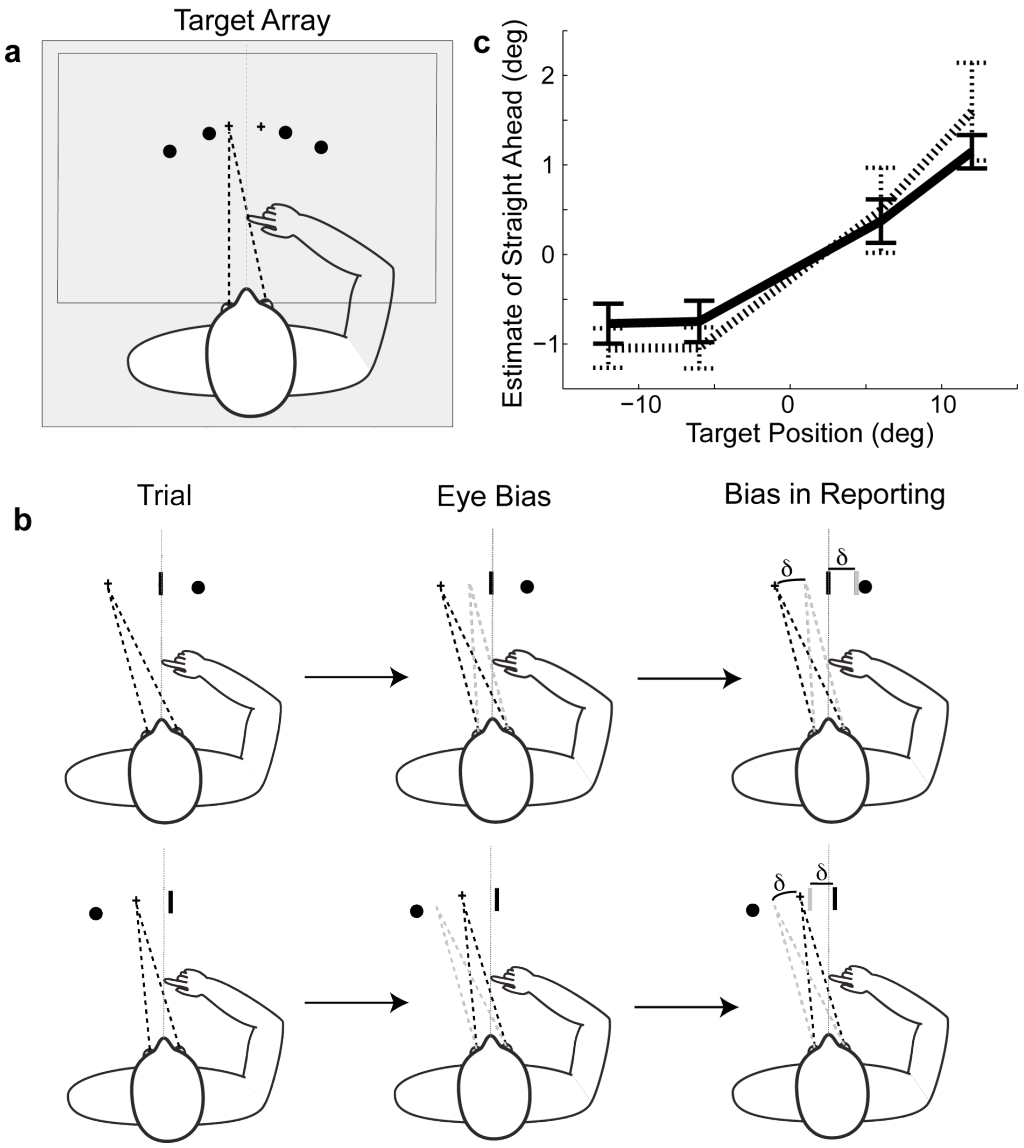


Figure S9: Inferring the eye bias via target-dependent bias on perceived location of straight ahead. a) Target array used for this experiment: + fixation points, • targets. b) Schematic of how bias in eye position toward the target would influence reported location of a transient visual stimulus relative to straight ahead. c) Inferred biases in the estimate of straight ahead. Solid line, no-reach trials; dashed line, reach trials. Error bars show standard error.

Chapter 2:

Mixed representations are shared across reaches to visual and proprioceptive targets in the superior parietal lobule.

Introduction

The simple tasks of reaching to pick up a coin or transferring the same coin from one hand to another without looking require similar movements, but use very different sources of sensory information. In the first case, vision of the coin enters the nervous system as a site of stimulation on the retina, so its location is defined in a retinotopic or eye-centered reference frame. In the second case, information about the position of the target hand enters the nervous system as mechanosensory signals about the state of the muscles and joints. These joint angles can be used to compare the position of the hand to the torso, a body-centered reference frame. These very different sensory signals are processed in the primary visual or somatosensory cortices, respectively, before converging in the parietal cortex, where they enter the motor circuit [3, 5-7, 102-106]. As the spatial information encoded by these two sensory streams must undergo costly [31, 37, 107] reference frame transformations to be compared [108], it is unclear what reference frame representation should be used when they converge in the parietal cortex.

A wide range of reference frames have been found in the superior parietal lobule (SPL) in studies of visually guided reaching [8-14]. Computational modeling suggests mixed or intermediate reference frames may be used to transform sensory information between reference frames [97, 109, 110]. Given the proposed role of the SPL in integrating sensory information arriving in a variety of reference frames for movement [3, 20, 33, 90, 111] this may explain the mixture of reference frames reported. However, there are consistent trends in the reference frame studies of the SPL, which suggest a link between the reference frames of sensory inputs to an area and the reference frames found in that area. For instance, while Area 5 responds to both visual and proprioceptive stimuli [21, 35, 90, 111-113], it receives greater somatosensory input [3, 103] and seems to have reference frame encoding that relates to hand or body centered (proprioceptive) reference frames [10, 12, 13]. Similarly, while MIP responds to both visual and proprioceptive stimuli [21, 90, 114], it receives greater inputs from visual areas in the occipital lobe [3, 103, 115], and seems to have reference frame encoding in eye-centered or eye-dependent (visual) reference frames [8, 11, 32]. Thus, one can ask whether the reference frames found in these areas are an inherent property of each area or are related to the reference frames of the sensory inputs.

Studies of multimodal cortical areas suggest that the reference frame of the sensory input affects, but does not completely determine, the reference frame of individual cells responses within a cortical area [16, 17, 27-29]. However, shared representations seem important for comparing and integrating sensory information [33, 38, 40, 116] (though see [27]). Further, a study looking at reaches to auditory targets in the parietal reach region

(PPR which overlaps MIP and V6A, [117]) found a similar representation to that used for reaches to visual targets [8, 15], though no within neuron comparison of responses to the two modalities was made. Thus, whether different sensory inputs are encoded in a common reference frame may depend on the role of the cortical area and task being performed.

Here we asked whether the sensory information available changes the reference frames used in the SPL. By directly comparing the reference frame representations used for reaches to visual and proprioceptive target we examined whether representations in Area 5 and MIP reflect the reference frame in which information enters the nervous system or is an invariant property of the cortical area. We found that the representations in these areas did not depend on the sensory information available, though a mixture of reference frame representations were present in both areas.

Methods

Experimental Setup

Two adult male rhesus macaque monkeys (12-15kg) were used in this experiment. All procedures were approved by the UCSF Institutional Animal Care and Use Committee and followed the NIH guidelines for care and treatment of laboratory animals.

The monkeys were trained to make reaches in a virtual reality setup allowing control of visual information during the task (Figure 1A). The monkeys were seated in a

primate chair with an open front panel open to allow arm movements. Head position was fixed with animals facing a mirror in which visual targets and feedback about hand position were presented. A digital video projector (NEC HT1100) displayed visual stimuli on a rear projection screen located directly above the mirror. The mirror and screen were positioned so that all visual objects appeared in the plane of the upper horizontal table on which the reaching arm rested. Eye position was monitored using an ISCAN (Woburn, MA) infrared eye tracking system. The monkeys were trained to wear a mesh jacket with stiff gloves that kept the hand prone. Magnetic sensors were attached to the gloves, and hand position was monitored with a Polhemus (Colchester, VT) Liberty magnetic tracking system. The arm contralateral to the recording chamber was used for reaching and rested on top of a thin (6 mm) horizontal table. The ipsilateral arm rested horizontally 5.5 cm below the upper table and was secured to a custom motor-driven sleigh that moved the arm passively between target locations. Behavioral and neural event times were recorded with a signal acquisition system that includes a programmable processor (Tucker Davis Technologies, Alachua, FL). Experiments were controlled with custom routines in Matlab (Natick, MA).

Target Modalities and Array

The monkeys were trained to reach to three different types of targets. They made reaches to visual (VIS), proprioceptive (the ipsilateral hand, PROP), and visual and proprioceptive (VIS+PROP) targets (Figure 1B). Visual targets were a presented as filled disks 2 cm in diameter. The disks were green during VIS trials and blue during

VIS+PROP trials to distinguish purely visual trials from trials were the target coincided with the position of the ipsilateral hand. The proprioceptive targets were the last joint of the two middle fingers of the ipsilateral hand, which were moved to the correct location with the positioning sleigh. These three trial types were preformed for the same set of reach conditions, i.e. reach target location, fixation point, and start location.

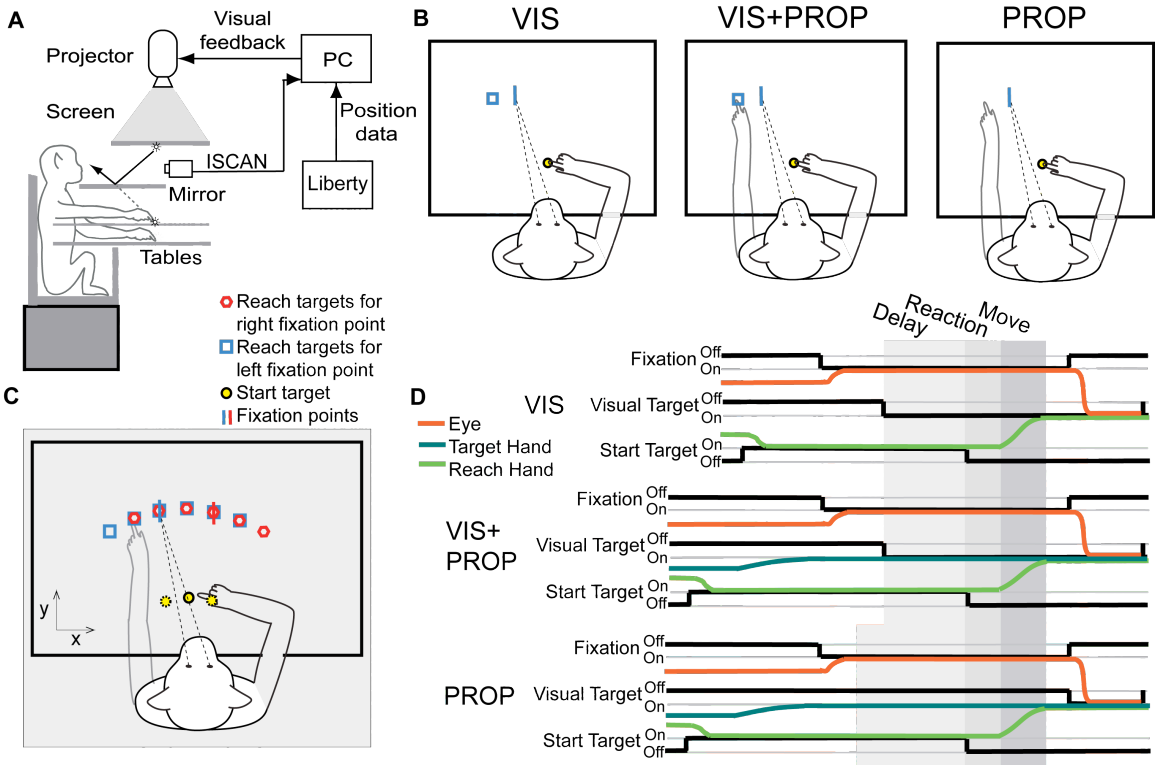


Figure 1: Experimental setup and time series. A) Schematic side view of experimental rig. B) Three target conditions. □ visual reach target, | fixation point, o feedback about reaching hand. C) Array of reach targets and fixation points. Only the center start location was used during most recording sessions, the additional right and left start positions were used for expanded start condition sessions. D) Schematic trial timeline for the three modalities. Colored lines show changes in hand and eye positions. Black lines show visual start, target and fixation point onset and offset. Grey shaded regions show behavioral epochs of interest.

Reach targets for all trial types were located in an arc equidistant from the projection of the monkeys' cyclopean eye position onto the tables (Figure 1A,C). This

point on the bottom table was also the pivot point for the lever arm of the positioning sleigh. The exact length of the arc was determined by the extension of the ipsilateral arm and varied slightly as the hand moved within the positioning sleigh (average radius 26 cm monkey C, 22 cm monkey E). Targets were positioned at 10 degree intervals with respect to the origin of the target arc (approximately 10 degrees visual angle) from -30 to +30 degrees from midline (Figure 1C).

During reaches the monkeys maintained fixation at one of two fixation points located at ±10 degrees from straight ahead. The fixation point was a filled red disk 8 mm in diameter. For each fixation point reaches were made to only six of the seven potential target locations (Figure 1C).

All reaches were made from a visual start location with initial feedback about the reaching hand. The start location was a green disk 2.4 cm in diameter and the feedback was a white disk 1 cm in diameter, positioned on the last joint of the two middle fingers. During most recording sessions a single start target was used. This start target was located on the midline (measured from target arc origin: 15 cm forward monkey C, 11 cm forward monkey E). In a subset of recording sessions two additional start targets were used to look at the effects of initial hand position on neural responses. These start targets were located at ±20 degrees visual angle at the same distance from the origin as the central start location. Reaches were made to a limited subset of reach targets from the additional start locations (-30, -10, +10 degree targets for -20 start target, and -10, +10, +30 degree targets for +20 degree start target).

Trial Presentation Order

The three trial types were presented in blocks so that all combinations of reach conditions were completed for PROP reaches, then VIS+PROP reaches, and finally VIS reaches before the next block (repetition) started. The trial types were kept separate within the blocks so that the animals knew what type of trial they were performing. All combinations of trial conditions (target position, fixation point, and start location) for each trial type were presented in random order within the trial type section of the block. Error trials were repeated at the end of the trial type section within the block. All trial conditions had to either be successfully completed or a maximum number of unsuccessful trials (typically 5) had to be reached before the next subsection would start. The animals generally completed trials successfully before the maximum error number was reached.

Trial and Reward Structure

In order to successfully complete a trial the monkeys had move their contralateral hand to the reach target without failing to complete any of the sequence of positional holds and delay periods enumerated here (Figure 1D). *1) Start target acquisition:* The monkey moved the hand into the start target disk and held the hand there for 500 ms. *2) Target hand positioning:* On PROP and VIS+PROP trials the ipsilateral hand was moved to the target location. *3) Fixation point acquisition:* The fixation target appeared and the monkey fixated it, maintaining fixation within 10-12 mm (~2.5 degrees visual angle) of

the initial fixation point in the X coordinate (in line with the target arc). *4) Visual target presentation:* After a 700 ms fixed delay the visual target appeared on the VIS and VIS+PROP trials. The same delay was used in PROP trials, though no visual target appeared. *5) Instructed delay:* The monkey maintained fixation and position at the start location for an additional variable delay of 500-1000 (Figure 1D, light grey). *6) Go-signal:* A go-tone sounded and the start target was extinguished indicating that the monkey should move the contralateral hand to the reach target (Figure 1D, light grey). *7) Reaction time:* The monkey began the reach after the go tone (Figure 1D, medium grey). When the hand moved 1 cm from the initial position, feedback of the hand was extinguished to eliminate the possibility of stimulating cells with visual motion. *8) Movement:* The monkey had to reach without stopping to a point within a set distance from the center of the reach target (Figure 1D, dark grey, monkey C: 4 cm VIS and VIS+PROP, 5 cm PROP; monkey E: 3 cm VIS and VIS+PROP, 4.5 cm PROP). *9) Target hold:* The final position had to be held for 200 ms to successfully complete a trial. *10) Reach feedback:* On successful trials the fixation point was extinguished and visual feedback for the contralateral hand was turned back on for 500 ms, providing visual feedback about endpoint distance from the reach target. On unsuccessful trials a 1 second error signal indicated which of the trial holds the monkey had violated. *11) Reward:* Monkey C received a water or fruit juice reward after successful trials. Monkey E received a food reward in the form of a slurry of monkey biscuits, apple juice, and banana after successful trials. Unsuccessful trials had no reward and a 1-5 second timeout before the next trial began.

To encourage accuracy, the amount of the reward depended in part on the proximity of the reach endpoint to the center of the reach target. Reaches within the inner third of the reach window received a full reward, while reaches outside this range received a reward that was half of the total reward scaled by the distance from the target divided by the size of the reach window. This reward never dropped below the minimum reward time of 50 ms (monkey C liquid reward delivered at ~1-1.2 ml/sec, monkey E slurry reward delivered at ~2.6-3.4 ml/sec). The maximum reward time was fixed within a block but increased at predetermined intervals throughout the day to keep the monkeys motivated. This reward scheme resulted in very similar rewards across trial types (Table 1).

A small portion of trials (10% in monkey C and 5% in monkey E) served as catch trials. In these trials the monkeys were rewarded without making a reach if they successfully maintained fixation through the final delay period (5), which was always extended to 1.5 seconds. These trials served the dual purpose of encouraging the animals to maintain fixation and preventing them from anticipating the go-signal on long delay trials.

Recording Cylinders

Both monkeys were trained extensively on the tasks before physiological recordings began. Immediately before the start of recording a 18 mm inner diameter titanium recoding cylinder, was positioned over a craniotomy opened above the IPS

(monkey C: 11 mm left, -4 mm posterior; monkey E: 12 mm right, -8 mm posterior, stereotactic coordinates relative to earbars). Monkey E had previously had chronic recording arrays implanted over his left motor and premotor cortices. The craniotomy was positioned with the guidance of structural Magnetic Resonance Images (MRIs). The animals were anesthetized and placed in a stereotax for the procedure. All surgical procedures and post operative care followed NIH guidelines.

For both monkeys, all recording sessions occurred within 10 months of implanting the recoding cylinder. Periodically, the dura mater was thinned to allow electrodes to penetrate into cortex. Mitomycin C, an antimitotic agent, was applied to the recoding chamber of monkey E to minimize tissue growth and reduce the required frequency of dural thinning.

Neural Recording

Single electrode recordings were used for all data collection. A Narishige Microdrive was used to lower a 2 mΩ (nominal) Alpha Omega tungsten electrode into cortex. All well isolated neurons that appeared modulated by the task were recorded without pre-selection for direction tuning. Neurons were recorded until the monkey completed 6-8 blocks (repetitions) or until isolation was lost. All neurons for which at least four blocks of trials were completed were included in further analysis. Included neurons had anywhere from 4-13 blocks with a mean of 6.2 blocks and a median of 6 blocks.

After recording, spikes were sorted and individual neurons were identified using Plexon Offline Sorter. Neurons were identified as having been nominally recorded in Area 5 or MIP based on the xy position of recording within the chamber and the depth of recording. The MRIs used for cylinder placement were used to make a map of the cylinder that, along with the observed neural responses to eye and hand movements, was used to identify neurons recorded from the SPL (Figure 2). Neurons recorded less than 2000 μm from the surface of the cortex in the SPL were categorized as Area 5 neurons while neurons recorded below this depth were categorized as MIP neurons. This depth cutoff was chosen because it typically corresponded to a quiet period, after the cells on the surface of the cortex, in which the electrode was likely passing through white matter.

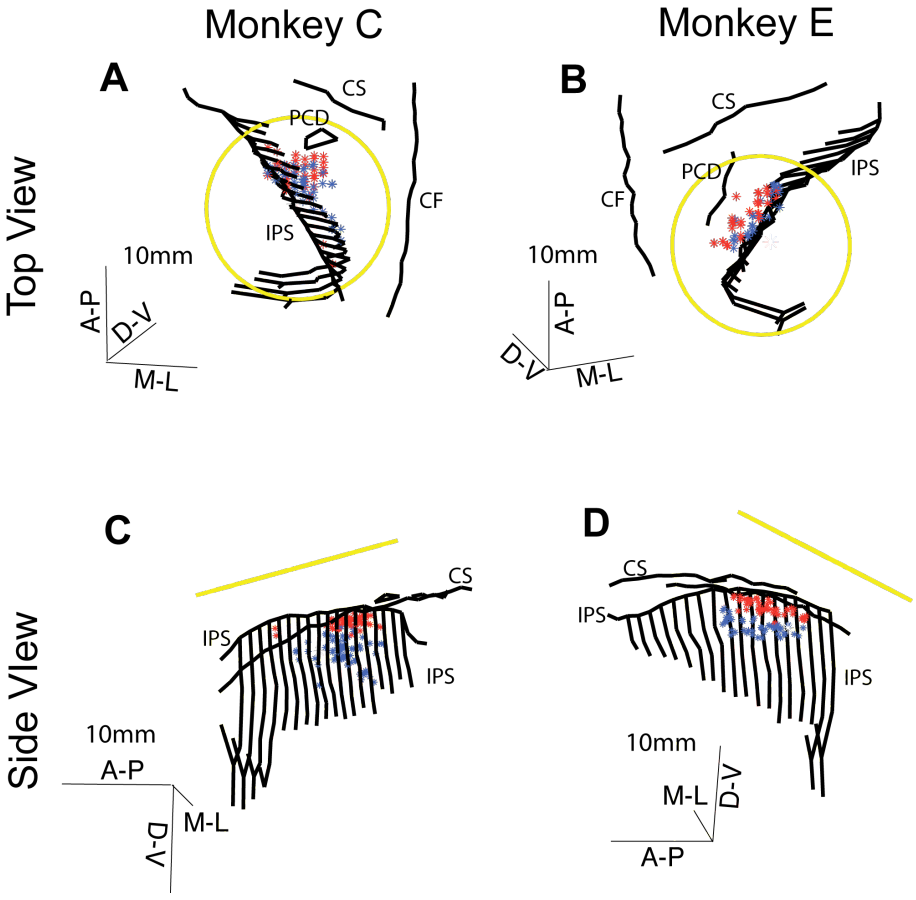


Figure 2: Inferred recording locations for monkeys C and E. Landmarks, black lines, are taken from structural MRIs used for cylinder placement. Parallel black lines show contours of the IPS. Yellow indicates cylinder position. Red stars show nominal Area 5 cells. Blue stars show nominal MIP cells

Behavioral Epochs For Analysis of Neural Data

Neural firing rates were averages across three different behavioral epochs defined as follows (Figure 1D, grey boxes). 1) The instructed delay (Delay) for VIS and VIS+PROP trials started at visual target onset and ended at the go-signal. Delay for PROP trials started at fixation point acquisition and ended at the go-signal. 2) The reaction time (Reaction) was started at the go-signal and ended at the initiation of movement. Movement initiation was the first time contralateral hand velocity exceeded 10 mm/s after the go-tone or the first time the hand moved more than 5 mm away from its start position. 3) The movement (Move) was from the start of movement to the end of movement. The end of movement was the first time the velocity of the contralateral hand dropped below 10 mm/s after movement initiation. The average firing rate on each trial during these epochs was determined for all neurons and used for the analysis described below.

Tuning Curve Fits: Reference Frame Shift and Gain Dependence on Eye Position

The firing rates of each neuron were fit separately to a tuning function for all target modalities and behavioral epochs. As our target array sampled only a portion of the workspace (60 degrees visual angle and ~85-110 degrees reach angle from start location), we chose a generic tuning function rather than the more traditional Gaussian tuning function. We found that a quadratic function (Equation 1) was generally sufficient to describe the cells tuning.

$$R = \alpha_1 + \alpha_2 (T - \delta \cdot E) + \alpha_3 (T - \delta \cdot E)^2 \qquad (1)$$

Here R is the firing rate on a given trial, T and E are the reach endpoint in degrees and fixation point in degrees, respectively, and the parameters α_1, α_2, α_3, and δ, where fit to the cells response using nonlinear least squares optimization (Matlab function lsqcurvefit). While using either target position or reach endpoint for T provided very similar fits, the reach endpoint fits generally had slightly higher R^2 values. Thus, all tuning curves reported here used reach endpoint as the target variable. The tuning function in Equation 1 includes a shift term (δ), which allows for the dependence of rate on reach endpoint (T) relative to the fixation point (E). This term was allowed to range from -0.5 to 1.5 and served as a smoothly varying metric for the reference frame for each cells response for a given epoch and modality. A shift value of 0 would correspond to a perfectly hand or body (the two cannot be distinguished with this analysis) centered representation while a shift value of 1 would correspond to a perfectly eye-centered representation. By bootstrapping the data across trials 1000 times [118] 95% confidence intervals and standard deviation for tuning curve parameters were determined. Cells had significantly different shift values across modalities when the 95% confidence intervals did not encompass the shift value of the other modality. In order to be included in the shift analysis of reference frame, cells had to have a 95% confidence interval for the shift (δ) term that was less than 75% of the allowable range of values. If one of the confidence intervals was pegged at the edge of the fitting range, the intervals were assumed to be symmetric and the range of this presumed confidence interval had to be less than 75% of

the fitting range. Further, cells had to have significant reach endpoint tuning ($p<0.05/9 = 0.0056$, Bonferroni correction, 3 modalities x 3 epochs), determined as described below, for inclusion in analysis.

Some cells showed gain modulation by eye position, thus, all cells were also fit to a tuning function which included an eye position dependent gain term instead of a shift term (Equation 2).

$$R = \lfloor \alpha_1 + \alpha_2 \cdot T + \alpha_3 \cdot T^2 \rfloor (1 - \gamma \cdot E) \qquad (2)$$

This gain term served as a secondary measure of eye position dependence of tuning. Some cells showed both shift and gain effects and so all cells were also fit to a tuning curve that included both shift (δ) and gain (γ) terms (Equation 3).

$$R = \lfloor \alpha_1 + \alpha_2 (T - \delta \cdot E) + \alpha_3 (T - \delta \cdot E)^2 \rfloor (1 - \gamma \cdot E) \qquad (3)$$

This model was selected to describe the cells response instead of Equation 1, to determine shift, or instead of Equation 2, to determine gain, if inclusion of the gain term (γ) or the shift term (δ) significantly improved the tuning curve fit in at least 2 of the 3 modalities in a given epoch. We used a single model selection criteria across modalities, despite the possibility of modality differences, because using the same model ensured that differences observed were due to differences in modalities and not the way different tuning cures were accounting for the same eye effects (gain and shift were often

redundant variables). Significant improvement in fit was determined using permutation tests with 1000 repetitions [93]. Briefly, for gain term inclusion, the sum square error (SSE) of the fit to Equation 3 was compared to the distribution of SSE values obtained from fits in which the values of E used in the gain term ($1-\gamma E$) were permuted across trials. Similarly, for shift term inclusion the SSE fit was compared to the distribution of SSE values obtained from fits to Equation 3 when the values of E used in the shift terms (δE) were permuted. The p-value for both tests was the percentage of times the permuted SSE was smaller then the SSE of the unpermuted fit.

Significant modulation of firing rate by reach endpoint was also determined using a permutation test with 1000 repetitions. Reach endpoints (T) in the tuning model selected (from Equations 1-3 as described above) were permuted with respect to firing rates and the p-value was the fraction of times the permuted SSE was smaller than the SSE with reach target information included.

Direct Rate Comparison of Reference Frames

It is possible that a poor tuning curve fit could result in miss-categorization of a cells reference frame. We therefore wished to examine the dependence of firing rate on fixation position without relying on tuning curve fits. In order to do this we looked at whether there were significant differences in firing rates between trial conditions that were matched in different coordinate frames. By looking at the category, or set of categories, in which firing rates were not significantly different we could determine the

reference frame that best described the cells response. We used a paired permutation test with 5000 repetitions to ask whether firing rates were significantly different (Bonferroni correction, $p < 0.05/3 = 0.0167$) between target pairs when targets were aligned in hand/body-centered coordinates (targets in same position relative to the body), eye-centered coordinates (targets aligned when plotted relative to fixation position, shifted over two targets in Figure 1C), or intermediate coordinates (in between hand/body-centered and eye-centered, shifted over one target in Figure 1C). For each modality and epoch a cell was categorized as responding in the reference frame (hand/body, intermediate, eye) or set of reference frames (hand/body or intermediate, intermediate or eye) for which there was no significant difference between paired firing rates. If the cells responses did not fit into any of these reference frame categories the firing rates were normalized for each fixation position and the test was repeated. This allowed cells that had gain modulation by eye position to be categorized by this analysis, and did not effect the pattern of reference frames observed.

Direct Rate Comparison of Target versus Movement Vector Representations

We used an analogous test to categorize the response properties of the cells recorded during reaches from multiple start targets. This subset of cells allowed us to look at whether responses where related to reach target position or movement vector, and the reference frame in which target and movement vector are encoded. We tested for significant differences in firing rate between reaches that had the same body-centered movement vector, the same eye-centered movement vector, or the same target position in

both body and eye coordinates but different movement vectors using a paired permutation test with 5000 repetitions. Cells were categorized as responding in the category or convergence of categories in which there was no significant difference between firing rates. As these categories tested very different aspects of the cells tuning, inferences could be draw about the tuning properties of the cells when responses were not significantly different than more than one category. For instance, if a cells responses were well matched when reaches were paired for eye-centered movement vector and target position independent of movement vector, then having the same eye-centered target position, was the key variable. The same argument holds for body-centered target coding when the responses are well matched in body-centered movement vector and target position independent of movement vector. This is because the cells' response does not depend on movement vector (as demonstrated by the target position independent of movement vector category), so only target position in a particular reference frame, eye or body, is important. Similarly if responses were well matched for both the body-centered and eye-centered movement vector then it is apparent that movement vector, which is hand-centered, is the key variable for that cell. If the firing rates were significantly different across all pairings, the firing rates for each fixation position where normalized to account for gain modulation and the test was repeated.

Results

Behavioral performance

Both monkeys were trained until performance on all three modalities plateaued before neural recordings began. Monkey C achieved a typical success rate of 60-75% correct trials a day. The low percentage of correct trials was due largely to a tendency to break fixation, particularly on visual reaches to the central target, which could not be reduced. Monkey E achieved a typical success rate of 90-97% correct trials a day. Both monkeys would typically perform 800 to 1400 total trials in a day.

The most striking behavioral difference between the three target modalities was the change in reach endpoint variance (Table 1). However, these changes are similar to those observed in humans using visual versus proprioceptive information for reaching [107] and reflect the differences in the reliability of the sensory modalities specifying target location [60, 70] not differences in the way the monkey is performing the task. This is supported by the fact that the VIS+PROP reaches have the smallest endpoint variability, suggesting that the monkeys were using both sources of sensory information about the target when available. Finally, the endpoint variance for all modalities was smaller than the distance between targets, so this variability does not affect our ability to determine the tuning for any of the modalities.

Behavioral parameters across modalities and monkeys.

		Peak Velocity (mm/s)	Reaction Time (ms)	Endpoint Variance (mm^2)	Reward (ms)
C	VIS	848 ± 218	366 ± 70	26.9 ± 5.4	425 ± 172
	VIS+PROP	831 ± 196	374 ± 77	23.0 ± 4.7	444 ± 166
	PROP	762 ± 179	418 ± 81	128.0 ± 35.2	437 ± 192
E	VIS	877 ± 133	361 ± 137	39.0 ± 9.1	323 ± 146
	VIS+PROP	892 ± 134	338 ± 110	33.7 ± 6.6	350 ± 135
	PROP	868 ± 130	354 ± 114	113 ± 27.9	337 ± 142

Table 1: Data for monkey C is from 63587 correct trials collected over 109 days. Data for monkey E is from 62672 correct trials collected over 70 days. Error is standard deviation.

Though velocity, reaction time and reward are similar across modalities, showing largely overlapping distributions, all differences shown are significant with $p<0.0001$ (t-test).

The monkeys' reaching movements were comparable across all three reach target modalities (VIS, VIS+PROP, and PROP). The animals performed smooth, rapid reaches with bell shaped velocity profiles in all trial types (data not shown). In order to quantify the relationship between movements for the three modalities we compared the peak velocity and the reaction time, two key parameters in reaching. While we found slight differences between the three modalities in these parameters, these differences are only a fraction of the standard deviation for these measures within any of the modalities (Table 1). Thus, the differences between the reaches performed for these modalities are small compared to the trial to trial variability in these reach parameters, and are unlikely to have large effects on the neural responses in parietal cortex. Similarly, the graded reward schedule successfully resulted in roughly equal rewards across the three modalities for each of the monkeys. This means that it is unlikely that differences in expected reward value played a role in neural responses in this study. All of these measures indicate that comparing neural responses to VIS, VIS+PROP, and PROP reaches will compare encoding of similar behaviors specified by different sensory parameters.

Cell tuning across areas and modalities

We recorded a total of 375 cells from two areas in the posterior parietal cortex: 193 cells from Area 5 (101 monkey C, 92 monkey E) and 182 cells from area MIP (95 monkey C, 87 monkey E) (Figure 2, and Methods). Of the cells recorded, 160 cells had

significant tuning in Area 5 (see Methods, 78 monkey C, 82 monkey E), and 164 cells had significant tuning in MIP (86 monkey C, 78 monkey E). Results were qualitatively the same across the two monkeys so all data is presented combined across animals.

Cells that met the tuning criteria (see Methods) were typically tuned for multiple modalities across epochs. Figure 3 illustrates this overlap in tuning with Venn diagrams. Within a given behavioral epoch there were slight differences in the proportion of cells tuned in the three modalities, but most of these differences were not significant ($p>0.4$, chi square test, Figure 3). The one exception is tuning during Delay in Area 5, which showed significantly fewer VIS tuned cells than PROP or VIS+PROP tuned cells ($p=0.001<0.05/6=0.008$, chi square test, Bonferroni correction). Many cells were tuned in more than one modality in a given epoch (Figure 3), though the proportion of cells tuned in all three modalities was not significantly different than expected by chance given the tuning for each modality ($p=0.059$ for Area 5, $p=0.396$ for MIP, chi square test). In both areas there were 40-102 cells tuned in each modality and epoch (modality-epoch) and a large percentage of cells were significantly tuned for each modality (68% in Area 5 and 82% in MIP had VIS tuning, 73% in Area 5 and 87% in MIP had VIS+PROP tuning, and 83% in Area 5 and 81% in MIP had PROP tuning). Additionally, many cells were tuned in more than one behavioral epoch (63% in Area 5 and 88% in MIP, see Figure 3 last column for overlap between epochs). The smaller proportion of cells tuned for more than one epoch in Area 5 may reflect the, on average, smaller target modulation seen in Area 5 compared to MIP (data not shown). However, the large number of cells tuned across

multiple epochs and modalities in both areas provided ample comparisons for examining differences neural responses to these modalities.

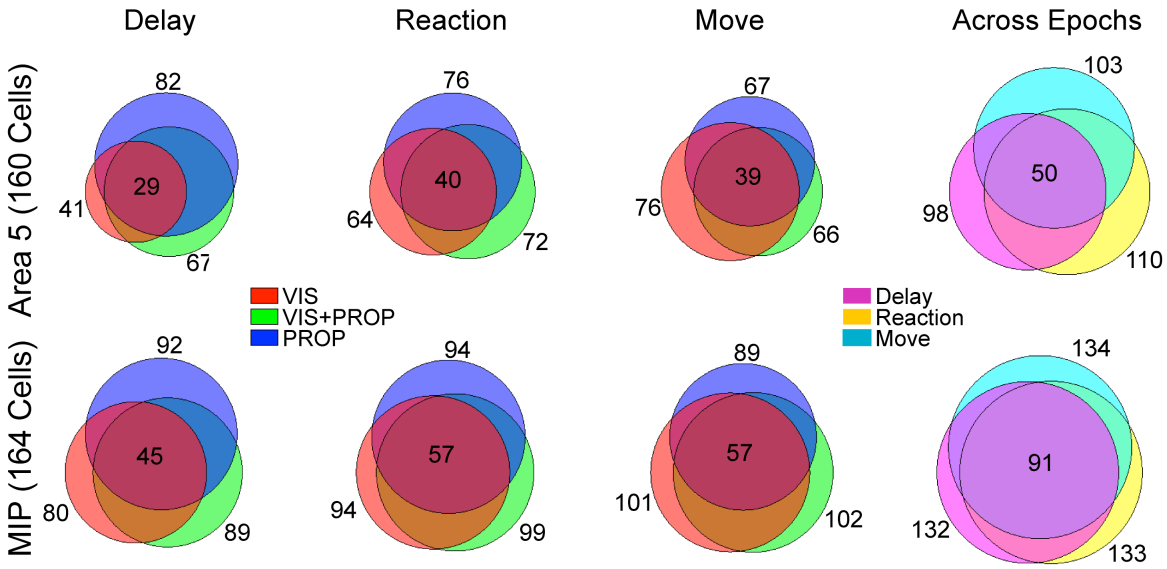


Figure 3: Venn diagrams showing number of cells with significant tuning in the three modalities across behavioral epochs. First three columns show number of cells tuned in each modality within that epoch. Last column shows number of cells tuned in any modality in the three behavioral epochs.

Reference frame effects on neural responses

We began by comparing the reference frames in which target information is represented across modalities. As reference frames within the parietal cortex have been the subject of debate [8-14, 16, 19, 20, 23, 38, 92, 119, 120] we used three different analyses to examine reference frame and eye position effects. We first determined the reference frame which best described the cells response in each modality using the reference frame shift analysis (see Methods) which allowed us to compare a continuous measure of reference frame

(measured as dependence of tuning on eye position) across modalities. We then compared responses across modalities when eye position effects were attributed to gain modulation rather than reference frame shifts. Finally, we used a direct rate comparison reference frame analysis (see Methods), similar in spirit to the reference frame analysis used to describe responses in these areas previously [8, 10, 15, 16, 23], to assign reference frames and compared these results to our continuous measure of reference frame. All of these analyses gave comparable results when modality, epoch, and area effects were examined.

Within cell comparison of tuning shift. Here we compare the reference frames that describe neurons responses during reaches to VIS, VIS+PROP, and PROP reaches. To do this, we fit a tuning curve to each modality that described the reference frame of the cells response in terms of a single parameter (tuning shift, δ, see Methods). The value of the tuning shift for each cell was determined by the dependence of tuning curve on eye position relative to reach endpoint. Our target and fixation point array (Figure 1C) was designed to distinguish between hand- or body-centered ($\delta = 0$) and eye-centered ($\delta = 1$) coding of reach endpoint. However, shift (δ) was a continuous variable, allowing us to capture intermediate reference frame responses, and was allowed to range between -0.5 and 1.5, so as not to over constrain the fits. As we will show below, the majority of fit values naturally fell between these boundaries.

Figure 4 shows and an example of this analysis for a single MIP neuron during Move. Note that the average firing rates for reaches made with the two fixation points do not align in either hand/body-centered coordinates (Figure 4A,C,E) or eye-centered

coordinates (Figure 4B,D,F). Indeed the ideal alignment between the two curves appears to be somewhere between these two extremes. The best-fit shift parameter for this cell ranges from δ = 0.64 in VIS to δ = 0.35 in PROP, corresponding to "intermediate" reference frame encoding. This analysis was performed for all tuned modalities and behavioral epochs for a given cell (all tuned modality-epoch-cell responses). All of the modality-epoch-cell shift fits that met the minimum criteria for having a best shift (see Methods) were used to look at the relationship between reference frame and modality, both within a cell and across the population in Area 5 and MIP.

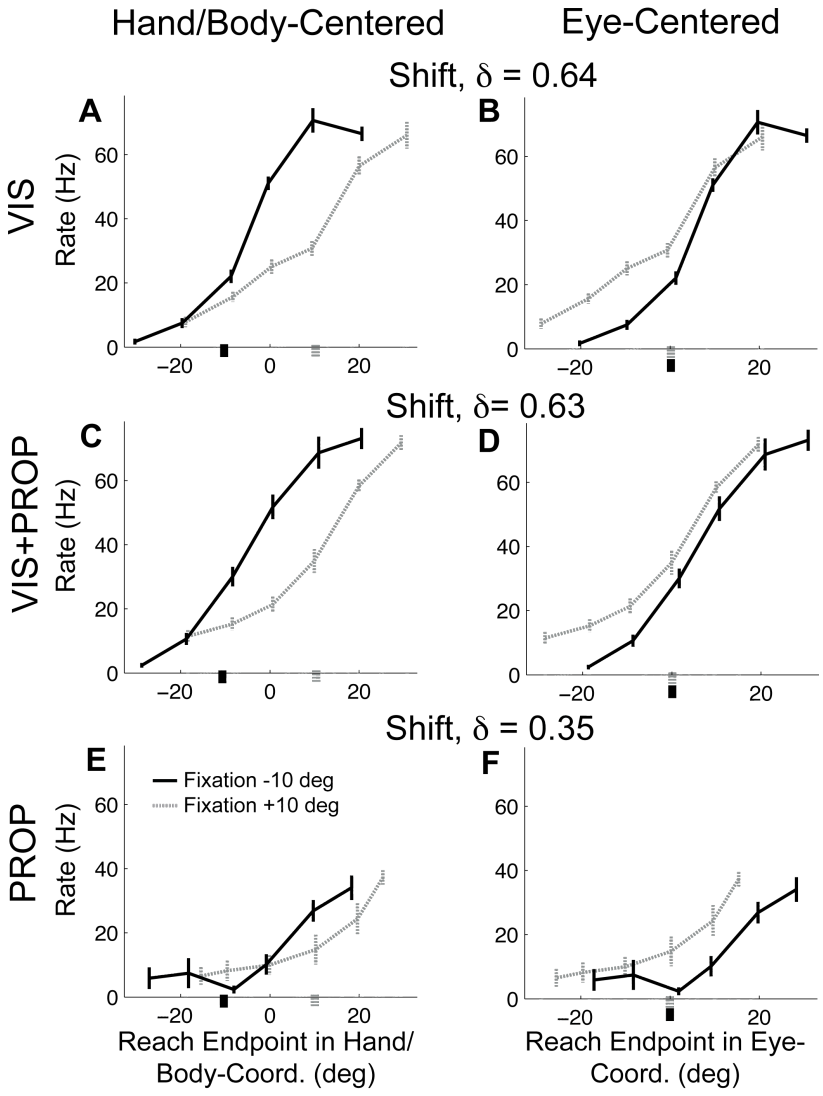


Figure 4: Example cell responses for the three modalities during Move. Panels show firing rate of the cell during reaches to the target array with responses for the two fixation points separated. Error bars show standard error. A,C,E) Responses aligned in body/hand-centered coordinates. B,D,F) Responses aligned in eye-centered coordinates. Shift values above each row show the fit tuning shift values for each trial type for this cell.

We first asked whether reference frame was modality dependent. To do this we compared the best shift (δ), across modalities for each epoch-cell (Figure 5, each epoch-cell is one point). The relationships between modalities' shifts were similar across behavioral epochs and cortical areas, though Area 5 and MIP showed differences in shift values that we will discuss below. Collapsing across epoch and area did not alter the results of this analysis or many of the analysis presented later, so Figure 5 and several later figures show comparisons for both areas and all behavioral epochs. The percentage of epoch-cells with significant differences across modality shifts (δ) depended on the modalities being compared (hollow versus filled points in figure 5, p=0.05 no correction for multiple comparisons, see methods). The smallest percentage of differences, 13.4 %, was between VIS and VIS+PROP (Figure 5A). The VIS+PROP and PROP comparison had an intermediate percentage of modality differences, 28.2% (Figure 5B). The largest percentage of differences, 33.3%, was seen between VIS and PROP (Figure 5C). These percentages were higher than would be expected by chance (5%) but the majority of cells for all three modality comparisons (66.7% for the weakest relationship, VIS and PROP) were not significantly different across modalities. These results suggest that tuning shifts are generally similar across modalities, thought not all cells have the same tuning shift across modalities.

Within these differences, one might expect the tuning shifts for VIS to be closer to 1 (more eye-centered) than the tuning shifts for PROP, since VIS target information enters the nervous system in an eye-centered reference frame (δ=1) and PROP target information enters the nervous system in a body centered reference frame (δ=0).

However, this is clearly not the case (Figure 5C, difference in mean p=0.543, paired permutation test, difference in distribution p=0.150, Kolmogrov-Smirnov test). Indeed, we found a strong correlation between the tuning shifts for any pair of modalities across epoch-cells (see Figure 5 for group R values, $p<0.001<0.05/3=0.017$ Spearman's correlation, Bonferroni correction). This relationship is strongest between the VIS and VIS+PROP shifts, followed by VIS+PROP and PROP, with VIS and PROP shifts showing slightly weaker, though still highly significant correlations. Thus, having common sensory information about target appears to strengthen the relationship between tuning shifts across modalities. However, the differences between VIS and PROP shift values within a cell and epoch do not appear to be related to the reference frames in which the different sensory information is naturally represented. Instead, these differences appear to reflect random deviations about the same mean representation.

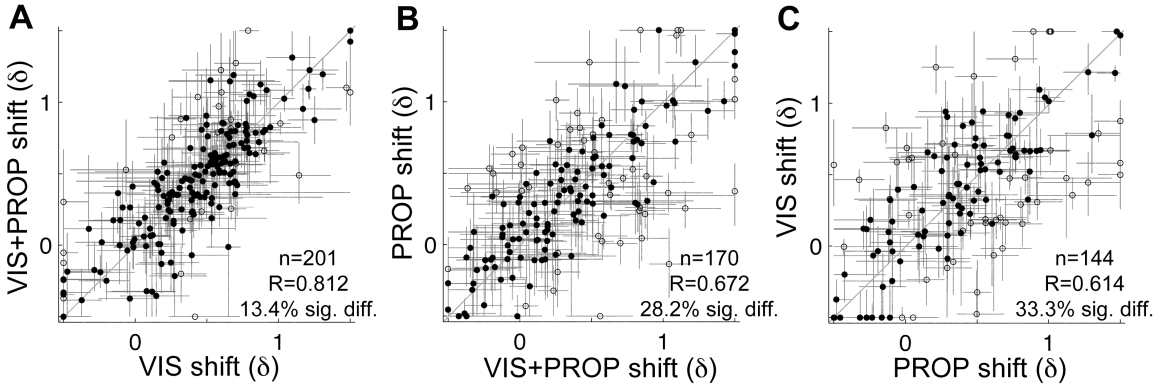


Figure 5: Comparison of tuning shift (δ) values across modalities for all epoch-cells, combined across Area 5 and MIP. All epoch-cells with best shift values (see Methods) in the pair of modalities being compared are shown. Error bars show bootstrapped standard deviation of tuning shift values. Correlations between tuning shift values across modalities were similar when responses were divided up by area (**Area 5:** VIS versus VIS+PROP R=0.773, VIS+PROP versus PROP R= 0.600, PROP versus VIS R=0.577, **MIP:** VIS versus VIS+PROP R=0.778, VIS+PROP versus PROP R= 0.637, PROP versus VIS R=0.533) and qualitatively the same when responses were divided into behavioral epochs.

Population distribution of tuning shift for modality, area, and epoch. In addition to within cell comparisons of tuning shift we examined tuning shifts for all modalities and epochs separately for Area 5 and MIP (Figure 6). This allowed us to look for differences in the tuning shift across the population that might only arise in cells that were tuned in only one of the modalities or epochs being compared. We found no significant differences between the mean ($p>0.3$, permutation test) or distribution ($p>0.2$, Kolmogorov-Smirnov) of tuning shifts for the three modalities in either Area 5 or MIP (Figure 6A&B). This again shows that there is no bias toward the reference frame of the sensory input in tuning shift values across modalities in these areas. There is, however, a slight trend in both areas toward higher shift values in later behavioral epochs relative to Delay (Figure 6C,D). This trend does not reach significance in Area 5 ($p=0.049>0.05/6=0.008$, Delay vs. Reaction, $p=0.017>0.008$, Delay vs. Move, permutation test, Bonferroni correction) but is significant in MIP ($p<0.001<0.008$ Delay vs. Move, permutation test, Bonferroni correction). This slight trend may be a real difference in encoding across epochs or may be related to the saccade the monkeys made to the reach target after the end of the movement (see Figure 1D and Methods), our data cannot distinguish these possibilities.

The most is a striking feature in Figure 6 is the difference between tuning shifts in Area 5 and MIP. This difference is highly significant between the mean ($p<0.001$, permutation test) and distribution ($p<0.001$, Kolmogorov-Smirnov test) of shifts across the two areas (compare Figure 6 A&C to B&D). Area 5 had a mean shift value of $\delta = 0.25$ while MIP had a mean shift value of $\delta = 0.51$ indicating that MIP had significantly

more eye-centered character in its coding of reach endpoint. These differences in shift values are consistent with differences in the information that can be read out from cells in these areas. When the cells from Area 5 and MIP were used to decode target position in either hand/body or eye-centered coordinates, we found that Area 5 contained more information in body-centered coordinates while MIP contained similar information in both reference frames (Supplemental Figure 1). However, it should be noted that both of these areas had mixed hand/body- and eye-centered coding, and there is considerable overlap in the distribution of shift values across the two areas. Thus, while the reference frame encoding of task parameters is clearly different across the two areas, neither area is well described by any single reference frame.

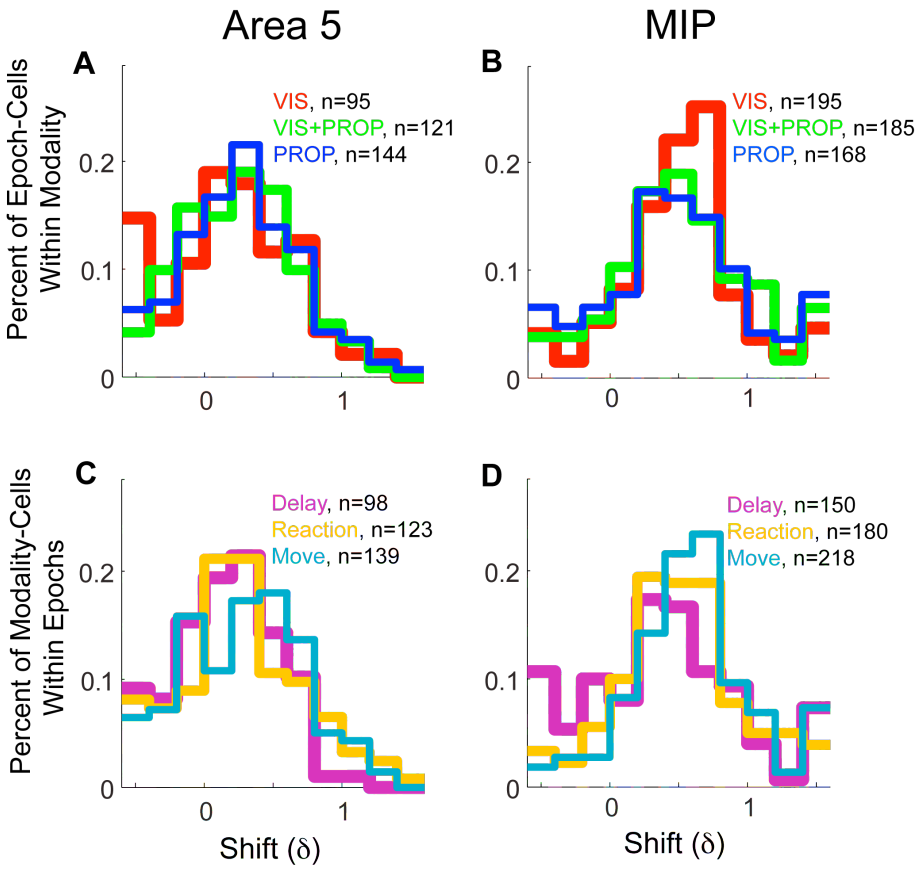


Figure 6: Histograms of tuning shift values in Area 5 (A&C) and MIP (B&D). A&B) Show the percentage of epoch-cells that had a given shift value within each modality. C&D) Show the percentage of modality-cells that had a given shift value within each epoch. The number of tuning shifts in each modality or epoch is shown in the figure legends.

While majority of modality-epoch-cells shown in Figures 5&6 have shift values which either lie between δ=0 and δ=1, or are not significantly different from these values, there are a few cases (122 of 1376 tuning shifts, 7.6% Area 5, 9.7% MIP) in which the fit tuning shift values differ significantly from this range. Many of these differences may be explained by noise, but there are a number of other explanations for these cells that should be examined. One possibility is that some of these cells may exhibit 'complex interactions' between eye-position and reach endpoint (see [16, 23]). Complex interactions mean that the cells response is not well described by any combination of 0-1 tuning shift and eye position gain effects. Such complex representations would push the shift values toward -0.5 or 1.5 as these values decrease the overlap between the two fixation-point tuning curves. Alternatively, a more conventional explanation for these extreme shift values is that these cells are actually gain modulated by eye position. Many of our cells are not sufficiently modulated across reach endpoints to distinguish a shift in reference frame from gain modulation by eye position. We examine the possible effects of gain modulation on cells responses in detail in the next section.

Additional analysis: gain modulation by eye position. The tuning shift analysis shown above favors reference frame shift over gain modulation as an explanation for eye position effects on tuning. This is a result of only including gain modulation by eye position in the tuning curve function if it significantly improved the model fit over a function with only tuning shift (see Methods). We wished to be sure we where not biasing our results by miss-categorizing responses, so we next carried out an analysis in

which we compared gain modulation across modalities in a tuning curve fit that only included a shift term if it significantly improved model fit (see Methods).

Describing eye position effects in terms of gain modulation showed similar relationships in eye position effects across modalities and areas to those seen in the shift analysis (Figure 7). This analysis included all cells and epochs that had significant tuning, unlike the shift analysis, because as cells that lacked best tuning shift values due to insufficient modulation still had a fit gain value. The inclusion of these cells did not change the relationship between eye position effects across modalities; the percentage of epoch-cell modality comparisons that are significantly different across modalities is nearly identical between the tuning shift and eye gain analyses, differing by at most 2.8% (compare percentages in Figures 6&8). Additionally, the correlations in gain values across modalities are nearly identical to the correlations in shift values across modalities, differing by at most 0.071. Thus, whether eye position effects are described in terms of gain or tuning shift there is a very strong relationship between the representations used across modalities for individual cells. The difference in eye position effects across areas is also significant when measured as either gain or shift. The distribution of gain values between Area 5 and MIP is significantly different ($p=0.001$, permutation test) with MIP showing a tendency to have larger absolute gain values (data not shown). Thus, the model implementation of eye dependent effects does not change the conclusions that representations are generally shared across modalities and differ between Area 5 and MIP.

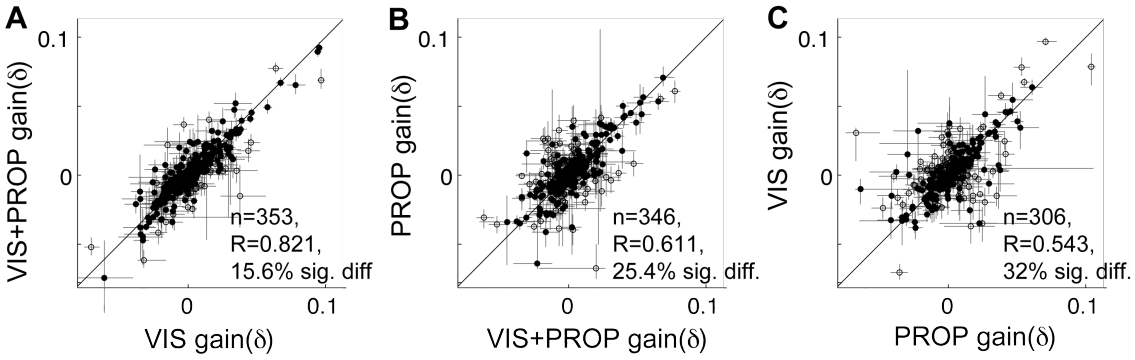


Figure 7: Comparison of eye-gain (γ) values across modalities for all epoch-cells, combined across Area 5 and MIP. All epoch-cells with significant tuning (see Methods) in the pair of modalities being compared are shown. Error bars show bootstrapped standard deviation of gain values. Correlations between gain values across modalities were qualitatively the same when responses were divided up by area or epoch.

Additional analysis: direct rate comparison of reference frames. Many researchers have looked at reference frame by comparing rates across trials which were paired in a given reference frame [8, 10, 12, 13, 15, 16, 23]. This type of analysis provides a measure of reference frame that does not rely on a tuning function fit. As a final test we compared the results of our tuning shift analysis to the reference frames assigned by direct firing rate comparisons in three discrete reference frames (hand/body-centered, shift = 0, intermediate, shift = 0.5, or eye-centered, shift = 1). Cells were assigned to the reference frame or pair of reference frames for which there was no significant difference between firing rates on spatially paired trials (see Methods). Many modality-epoch-cells had a best reference frame determined with both the tuning shift and direct rate comparison analysis. The reference frames assigned with these two analyses were in good agreement (Figure 8). Additionally, only 6.8% of modality-epoch-cells reference frames were significantly different than all three reference frame categories tested. Of these, 34% (of the modality-epoch-cells that also had a best tuning shift) had tuning shifts

that were significantly outside the zero to one range, suggesting that tuning for these cells lies outside the continuum of body/hand to eye centered reference frame encoding. However, the small percentage of cells that are significantly poorly described by any of the reference frame categories suggests that most cells responses can be described as a linear combination hand/body and eye-centered coding, such as our tuning shift model.

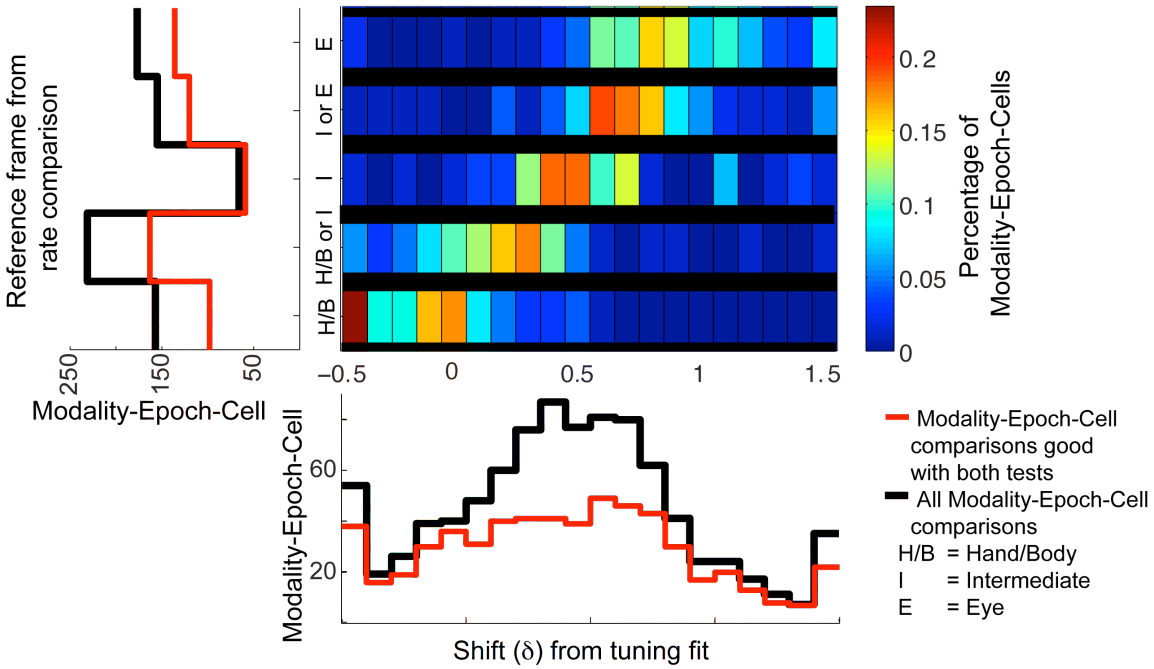


Figure 8: Comparison of tuning shift and direct rate comparison reference frames. Heat map shows percentage of modality-epoch-cell in a given direct rate comparison reference frame category that had a particular shift value. Histogram on y-axis shows distribution of reference frames from direct rate comparison analysis. Histogram on x-axis shows distribution of reference frames from tuning shift analysis. Black lines show all modality-epoch-cells with best references frames in each analysis (see Methods). Red lines show modality-epoch-cells that had best reference frames in both analysis and are included in the heat map.

While the reference frame assignments were similar across the two analyses the patterns of values assigned for each analysis (see histograms on X and Y axes in Figure 8) reveal that there were differences in the distribution of reference frames assigned with each analysis. The direct rate comparison analysis in particular has biases in assignments. This is due to the fact that there were more target pairs in the intermediate reference

frame (more chances for differences, making this category less likely), and the fact that inclusion in joint reference frames (e.g. intermediate/eye) only required rejection from one category while inclusion in the a single reference frame category required rejection from two categories (see Methods). Nevertheless, both analyses show similar trends in the relationship between reference frame and modality, behavioral epoch, and cortical area (data not shown). Like the tuning shift analysis the direct rate analysis had no significant difference in reference frame across modalities ($p=0.809$, Kruskal-Wallis test), though there were significant differences in reference frame across behavioral epochs ($p<0.001<0.05/3=0.017$, Kruskal-Wallis test, Bonferroni correction) and between Area 5 and MIP ($p<0.001<0.017$, Kruskal-Wallis test, Bonferroni correction). Thus, the main findings of this study do not depend on the method used to determine reference frame.

General conclusions from eye position effects analysis. The main conclusions from the three analyses of reference frame (or more generally eye position effects) on neural responses presented above are the same. First, and most importantly, the reference frame in which a cell encodes movement parameters appears to be independent of the sensory information used to specify the target, in both Area 5 and MIP. Second, there is a slight trend towards having larger eye position effects in later task epochs. This effect may represent a general trend across these areas or be the result of the animal planning to saccade to the reach target after the movement. Whatever the underlying cause of this trend, it does not appear to affect the relationship between representations across modalities and is relatively small in magnitude. Third and finally, there is a significant difference in the magnitude of eye position effects between Area 5 and MIP. However, it

must be noted that both areas exhibit a broad distribution of eye position effects and the majority of cells are not well described by a pure hand/body or eye-centered reference frame.

Tuning curve shape across modalities

Given the similarity in eye position effects across modalities, we wanted to ask the more general question of how similar tuning curve shape is across modalities. To do this we looked at the correlation between firing rates for all trial conditions across modalities. This provided a quantitative measure of the similarity in the shape and preferred direction of tuning across modalities. Both areas tended to have high positive correlations between modalities for an epoch-cell (well above the correlations seen when pairs of epoch-cells were chosen randomly), indicating that tuning was very similar across modalities in both Area 5 and MIP (Figure 9A-C).

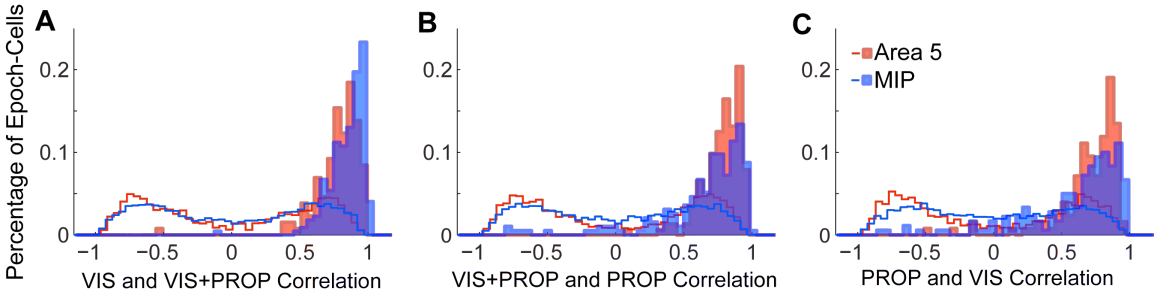


Figure 9: Histograms of correlation coefficients between firing rates in all reach conditions across modalities. Filled histograms show percentage of epoch-cells that had a given correlation coefficient the pair of modalities being compared. Hollow histograms show the percentage of times a given correlation coefficient occurs across the population when pairs of cells are chosen at random.

In both areas the correlation between firing rates across modalities was weakly negatively correlated with differences in tuning shift (R values range from -0.236 to -0.285, $p<0.002<0.5/3=0.017$, Spearman's correlation, Bonferroni correction). This indicates that epoch-cell responses with larger differences in reference frame across modalities had lower correlation between modality tuning curves, as would be expected. Overall these results suggest that the similarity in reference frames across modalities in these areas is related to a more general correlation in tuning properties across modalities.

Encoding multiple movement parameters

The analyses presented so far have looked at cells tuning with respect to target or reach endpoint and eye position in terms of the reference frame of encoding. We found that tuning properties of cells with respect to these parameters are largely invariant across modalities. However, there are additional movement variables, namely the position of the reaching hand at the start of movement, and the relative positions of the reaching hand and target or eyes, that have been shown to be encoded by the parietal cortex [10, 12, 13]. Here we consider how these parameters affect the responses of neurons in Area 5 and MIP.

By adding additional start locations to our task conditions for a subset of cells, we were able to look at how target and movement vector encoding interact with eye position or reference frame effects (see Methods). We collected extended data sets for 87 Area 5 cells (20 monkey C, 67 monkey E) and 57 MIP cells (12 monkey C, 45 monkey E). Of these cells 62 Area 5 cells and 51 MIP cells were significantly tuned ($p<0.05/9=0.006$,

ANOVA, Bonferroni correction). The fact that the additional start position data was collected at the same time as the full dataset used to determine reference frame limited the number of reaches that could be added, limiting the power of this analysis. However, we were able to add sufficient differences in start location and movement vector to place many modality-epoch-cell responses into one of six possible reference frames (see Methods for details, Figure 10A). Three of these reference frames encoded the reach movement vector in body-centered, eye-centered, or hand-centered coordinates (pure movement vector coding). The remaining three reference frames encoded target position independent of initial hand position in body-centered, eye-centered, or mixed eye- and body-coordinates (responses were well described when targets were matched in both eye- and body-coordinates but not in either reference frame alone). A small portion of cells, 4.9%, had a significantly poor fit in all possible categories. The existence of such cells is not surprising as the reference frames examined were very regimented and our shift analysis suggests that many cells exhibit intermediate reference frame coding. Nevertheless, modality-epoch-cell responses for which we were able determine a best match reveal distinct trends within the data.

The results of this analysis are consistent with the reference-frame analyses presented above in that they show no significant modality differences and only weak epoch differences in reference frame. When the epoch-cell responses are compared across modalities (Figure 11B) there are no significant differences across the six reference frame categories listed above ($p=0.962$, chi-square test). The similarity in modality responses holds even when these categories are grouped to compare movement vector and target

coding (first three categories versus last three categories in Figure 10B, p=0.301, chi-square test). There are no significant differences in between epochs in the six reference frame groupings (Figure 10C, p=0.330, chi-square test). However, there is a transition from having more movement vector coding in Delay to having more target coding in Move, though this is not quite significant, even when movement vector and target groupings are directly compared (p=0.009>0.05/6=0.008, Chi-square test, Bonferroni correction). As with the epoch differences observed above, this may reflect a gradual change in the encoding properties of the cell over the course of the movement or planning for the saccade made after completion of the reach.

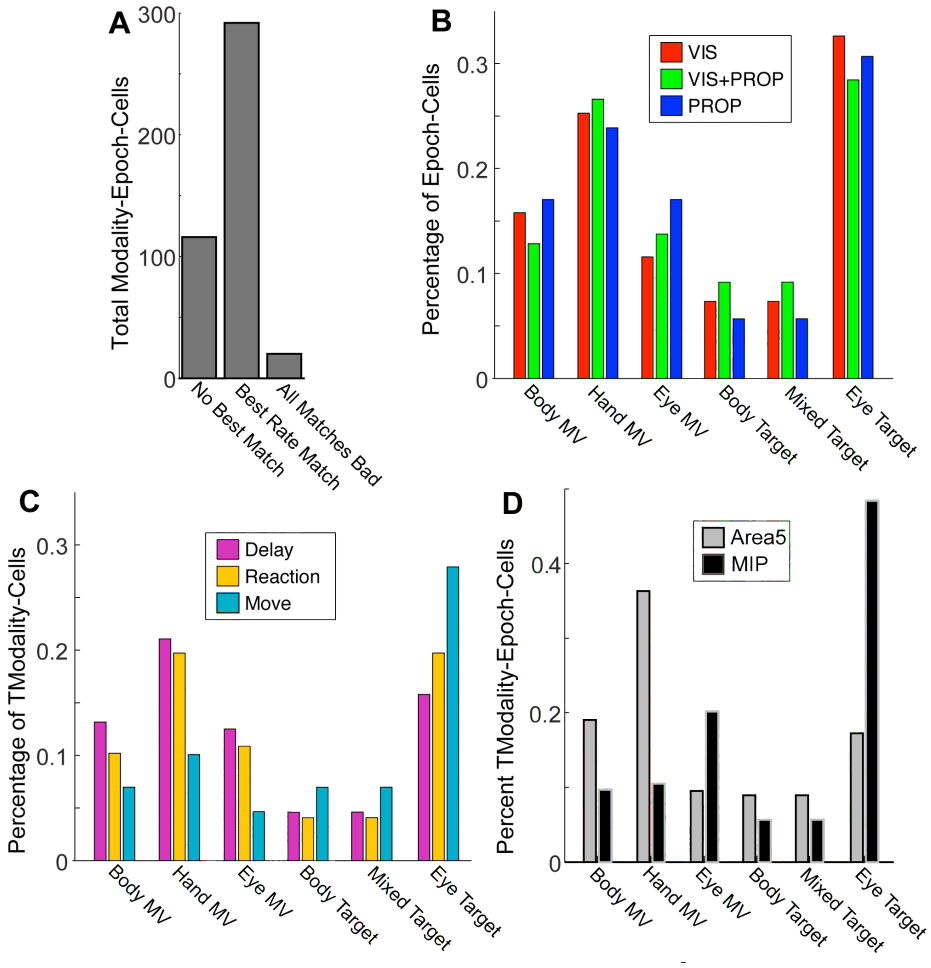


Figure10: Target versus movement vector reference frame responses of cells recorded during additional start location sessions. A) The number of modality-epoch-cells that were not

significantly different when trials were paired in any of the reference frames tests, had a best reference frame of the reference frames tested, and were significantly different than all of the reference frames tested. B) The percentage of epoch-cells within a modality that fell into each reference frame category. C) The percentage of modality-cells within an epoch that fell into each reference frame category. D) The percentage of modality-epoch-cells in Area 5 and MIP that fell into each reference frame category. The first three reference frame categories in B-D are different reference frames with movement vector coding of reach responses. The last three reference frame categories in B-D are different reference frames with target position coding of reach responses.

The most striking difference in reference frame categories is between cortical areas (Figure 10D). While both Area 5 and MIP have conditions that fall into all six possible reference frames, Area 5 has a distinct peak in hand-centered movement vector coding, while MIP has a distinct peak in eye-centered target coding. The differences between these areas are significant for both the fine scale division of reference frame and the movement vector versus target comparison ($p<0.001<0.008$, Chi-square test, Bonferroni correction). The distribution of responses across Area 5 and MIP are consistent with the reference frame analyses presented above in that both areas exhibit a mixture of reference frames, with Area 5 showing more hand/body-centered coding, and MIP showing more mixed and eye-centered coding. Even with the rough categorizations used in Figure 10, both areas show a range of coordinate frame representations. Thus, any attempt to define the coding of these areas in a single reference frame will miss important properties of the population responses.

Discussion

The main purpose of this study was to investigate whether reference frames in the SPL used for reaching depended on the sensory modality used to specify target location.

There is evidence from other cortical areas suggesting that reference frame representations are often affected by the sensory modality of the input being represented [16, 17, 27-29, 121]. To investigate this possibility, we compared reaches to visual targets and proprioceptive targets (the ipsilateral hand) in two parietal areas, nominally Area 5 and MIP (see Figure 2 for recording locations). As Area 5 and MIP are involved in early integration of these sensory inputs for movement planning, changes in representations related to the sensory information available would likely be apparent in these areas. We found that while reach representations differ between these two cortical areas, they do not appear to depend on the sensory modality used to specify the movement.

The reference frames in movement parameters are represented in the parietal cortex have been the subject of much debate [8-14, 16, 19, 20, 23, 38, 92, 119]. We found that both Area 5 and MIP exhibited a mixture of encoding schemes, with many neurons showing intermediate reference frame representations. Within the observed distribution of responses, Area 5 neurons had a slight preference for movement vector encoding, and a trend towards more hand/body-centered representations (Figure 6A&C and Figure 10D). MIP neurons showed a slight preference for target position coding, with trend towards intermediate reference frame representations (Figure 6B&D and Figure 10D). Our results are broadly consistent with previous studies that found that Area 5 exhibits primarily body-centered or hand-centered [12, 13] coding while MIP and PRR exhibit primarily mixed eye and hand [11] or mixed eye and head/body-centered coding [16, 23]. However, our findings do not, at first, appear consistent previous studies showing that Area 5 encodes eye-centered movement vector [10] and PRR (which overlaps MIP) encodes eye-centered

target position [8]. Part of the discrepancy between our findings and those of Batista et. al. may be explained by recording location, as we were likely recording anterior to PRR [117]. However much of the apparent discrepancy between our results is likely due to these studies overly regimented characterization of neuron responses into a single best reference frame response. This is illustrated by the recent finding that when reference frames in PRR are measured with a continuous measure of reference frame, similar to the shift value used in this study, a wide distribution of responses in seen, with a peak near the eye-centered coding scheme previously described [8, 11]. Taken in this context, our results provide a more detailed, but largely consistent, picture of the encoding scheme found in the parietal cortex. Further, our results combined with the results of Chang et. al. suggest a continuous gradient in the distribution of reference frames across the SPL, with broadly eye centered coding in the posterior IPS [11], intermediate eye- and body coding further forward in the IPS (MIP) and broadly hand/body-centered coding on the surface of the SPL (Area 5).

There is an increasing body of evidence that different kinds of sensory information, arriving in distinct reference frames, are represented in a variety of intermediate and mixed reference frames throughout the motor system [4, 11-13, 16, 18-26]. Computational modeling studies have shown that mixed representations may arise naturally in networks of neurons performing coordinate frame transformations [97, 109, 110] and that pure representations would not necessary be expected in areas encoding spatial information from multiple sensory modalities [27, 97]. Further, our decode model shows that variables can be reliably interpreted from these mixed representations in a number of

different reference frames (Supplemental Figure 1). This is not to say that the spatial information contained across different reference frames is not important. We have previously shown that there may in fact be some computational benefit, in terms of reduced variability, in having multiple reference frame representations for movement planning [107]. However, it is becoming increasingly apparent that while the idea of simple discrete reference frame representations within the cortex is attractive in its simplicity, it is not necessary or perhaps even desirable to encode movement variables in a series of discrete and pure reference frame representations.

The precise mixture of representations found in our study varied slightly over the course of movement planning and execution. There were slight trends in both Area 5 and MIP towards having more eye-centered coding (higher shift values, Figure 6C&D) and more target coding (or less movement vector coding, Figure 10C) in later behavioral epochs. It is impossible to determine from our study whether these differences represent real changes in the variables encoded over the course of the reaching movement or relate to the saccade the animal was allowed to make to the target 200 ms after the end of movement. A previous study looking at reference frame over the course of the reach did not find significant differences over time [9]. However, they only roughly categorized reference frame and did not examine time points as far into the movement as we did, so the subtle differences we found may have been overlooked. These changes could reflect fact that the initial hand position is important for planning a movement vector but is less relevant as the hand approaches the target. Alternately, the changes could reflect the change from movement planning to feedback control of the movement [38, 122-125]. Whatever

the underlying cause, the differences seen across epochs were very small relative to distribution of reference frame representations seen in either Area 5 or MIP across all modality types.

In examining the relationship between representations across modalities it is important to note that both motor and sensory responses are common in the parietal cortex [7, 21, 35, 90, 92, 111, 112, 114, 117, 126, 127] We chose to record from cortex contralateral to the reaching hand because contralateral firing rates are more closely related to movement parameters [128]. However, the parietal cortex has significant responses to the ipsilateral hand [112-114, 127, 128] as well as responses to the presence of the visual target [90, 112, 126]. Having these different sensory representations present should, if anything, increase the likelihood of seeing differences between representations used for visual and proprioceptive reaches. Indeed, this may explain the slight decrease in correlation between VIS and PROP reference frames relative the correlation between VIS or PROP and the VIS+PROP reference frames (Figure 5). However, there was still a strong relationship between reference frames for VIS and PROP reaches and no difference in the average coordinate frame representation for any modality comparisons.

The similarity in reference frames and tuning curves seen across modalities in this study is difficult to reconcile with the results of multimodal studies in other areas. Studies comparing the reference frames used to represent discrete sensory inputs have most commonly found reference frames are skewed towards the reference frame in which the sensory information in entering the nervous system [16, 17, 27-29]. These results suggest that there is no general requirement for the alignment of responses across the different target sensory modalities observed here. Some of the differences between our result and the

results of previous studies may be explained by differences in areas sensory versus motor roles. We chose to record ipsilateral to the target hand and contralateral to the reaching hand, which may have biased our results toward encoding of motor rather than sensory parameters [112, 114, 127, 128]. As the movements were similar across modalities (Table 1), this would increase the likelihood of seeing similar representations. However, given the bilateral responses in parietal cortex [112-114, 127, 128] this is unlikely to fully explain our results. Indeed, several studies that showed differences in reference frame across target modalities for saccades were in LIP and SC [16, 17, 29] which are important for the targeting and execution of saccades [129] similar to the SPL's role in reaching [3, 20, 38]. Further, studies in some of the same areas, using different behavioral tasks, have found no difference in reference frame representations across sensory modalities [15, 130] similar to the results of this study. Thus, the most important element in understanding why we see similar reference frames across different target modalities may be the role of Area 5 and MIP in using visual and proprioceptive information for reaching.

The SPL is thought to function in the representation of spatial information for movements across multiple effectors [19, 90, 92, 126] and in the integration of sensory information [35, 112, 131-133], both of which would be simplified by sharing common representations [33, 92, 116, 131]. Our finding that modality representations are similar could be related to the similar tuning seen across tasks in the posterior SPL [90, 126]. This 'global tuning' is hypothesized to reflect a strategy for coordinating the movement of eyes and hands in space [19, 92, 134, 135]. In this case, movements involving the two hands would simply be another kind of effector interaction to bring into spatial register. The SPL has also

been suggested to be important for maintaining a unified position estimate for vision and proprioception of the hand [34] and in integrating this information [35]. In daily activities, both visual and proprioceptive signals about hand position are generally available and congruent, and sharing similar tuning across these inputs may be important for integrating this information for movement planning. Indeed, having the same set of representations independent of sensory modality of the target may be beneficial when integrating sensory information [107]. In support of this idea, a recent study in MST showed that cells with similar representations across sensory inputs are more likely to be drawn upon during sensory integration [136]. These results suggest that the correlation between visual and proprioceptive representations could be the related these areas' role in maintaining unified spatial percepts [34, 92, 135] and integrating sensory information [35] for movement planning [3, 4, 21, 137, 138].

Grants

This work was supported by the National Eye Institute (R01 EY-015679), the National Institute of Mental Health (P50 MH77970). L.M.M.M. was supported by a graduate fellowship from the National Science Foundation and an NIH Vision Training Grant.

Supplemental Information for Chapter 2: Mixed representations are shared across reaches to visual and proprioceptive targets in the superior parietal lobule.

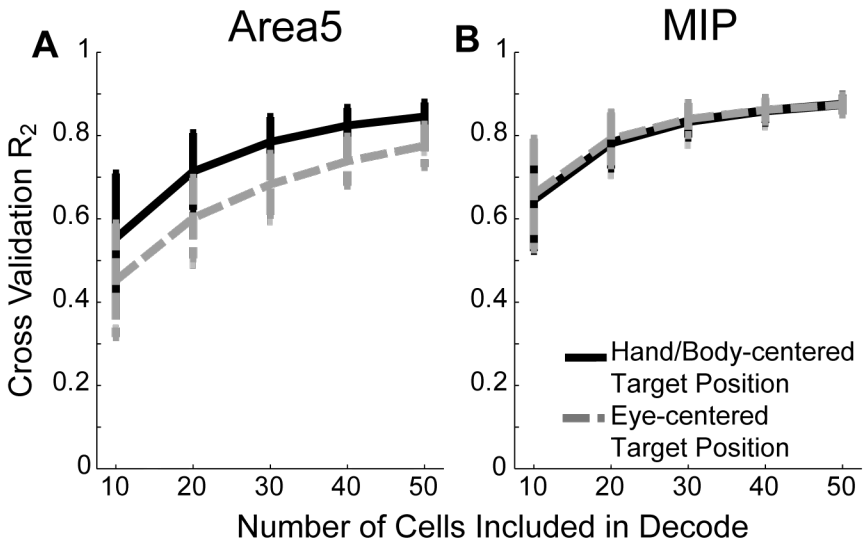
Decoding movement parameters from mixed representations

Given the heterogeneous reference frames observed in both Area 5 and MIP, we set out to determine whether the observed differences in the distribution of reference frames across these areas had a meaningful affect on the information that can be decoded from them. In order to do this we built a linear decoding model. We used linear regression with the desired readout (hand/body-centered target position, or eye-centered target position) as the dependent variable. Because we used cells recorded across many recording sessions for this population analysis, we had to use target position rather than reach endpoint for this regression. The individual trial firing rates of a subset of cells were used to generate the independent variables for the regression. All cells reported in this paper had at least four blocks of repetitions, though the majority of cells had more. In this analysis we discarded any repetitions beyond the first four so that all cells would have the same number of trials. Because the firing rates of the neurons are correlated when they have similar tuning curves, we first computed the principle components of the

firing rates using Principle Components Analysis (PCA) and projected the cells firing rates onto the principle components. The projections of the rates where then ranked according to how much of the dependent variable response they explained. The projections where added into the regression one at a time in order of their contribution and a leave out one trial out cross validation R^2 value was computed. Rate projections were added to the regression until the cross validation R^2 failed to decrease by more than 1% of the previous R^2 value. This process was carried out for 1000 random subsets of neurons containing 10, 20, 30, 40, and 50 cells for each epoch and modality in both areas for both monkeys. From these measures we computed a mean and standard deviation for the cross validation R^2 for each possible readout and sample size. This model was not intended to mimic the brain, it was meant simply to provide us with a measure of the information encoded by the firing rates of a group of cells.

The results of this decode analysis revealed a number of distinct trends in the population decodes across areas. In both Area 5 and MIP the R^2 values rapidly asymptoted for most readouts as the number of cells included in the decode increased (Supplemental Figure 1). In Area 5 the highest R^2 decode values were obtained for the hand/body-centered target location with eye-centered target location having the second highest value (Supplemental Figure 1A). This indicates that Area 5 encodes more information in hand/body-centered coordinates than eye-centered coordinates, which is consistent with the reference frame analyses presented in the main text. It is important to note that the target position in hand/body and eye-centered coordinates are highly correlated variables, so the fact that the R^2 values were still relatively high in the Area 5

eye-centered readout is not surprising. MIP on the other hand, had almost identical R^2 values for the hand/body and eye-centered decodes (Supplemental Figure 1B), indicating that there is roughly equal amounts of information about these two variables in the population response of MIP. This is consistent with the average tuning shift value of δ = 0.51 for MIP cells. Thus, the differences in reference frame distributions across areas are related to the fidelity with which movement parameters in different reference frames can be decoded from these areas.



Supplemental Figure 1: Population decode of reach target in eye-centered coordinates, reach target in hand/body-centered coordinates, and eye position, with different population sizes for Area5 and MIP. Error bars show standard deviation of decode R^2 value across 1000 randomly selected groups of cells (see Methods).

Chapter 3:

Integration of visual and proprioceptive information in population responses of the posterior parietal cortex.

Introduction

When reaching to swat a mosquito on ones arm, a person may rely exclusively on somatosensory information to target the movement, or may also look at the arm to gather additional sensory information and improve his chances of squashing the mosquito. Making use of and integrating multiple sources of sensory information are a fundamental part of animals' interactions with their environment. Psychophysical studies have shown that humans integrate sensory information in a statistically optimal manner under a variety of conditions [65, 66, 70, 139]. However, while many studies have investigated how individual neurons change their responses when multiple sensory inputs are present [116, 120, 136, 140-145] little is known about how this integration is carried out in multimodal areas of the brain.

A major focus of previous studies of sensory integration has been the responses of individual neurons in the superior colliculus (SC), which is important for saccadic eye movements [129]. These studies found that the firing rates of many cells change when multiple sources of sensory information are available [116, 140-142]. Super-additive bimodal responses, responses that are greater than the sum of the unimodal responses, have received much attention [140, 142]. However, many factors, such as spatial and temporal coincidence of stimuli, and stimuli strength, affect how cell responses change with bimodal stimuli, and additive responses are much more common when strong sensory stimuli are used [116, 141, 142]. Additive responses to bimodal stimuli match the predictions of network models of sensory integration, which demonstrate that the enhancement of individual cells bimodal responses, embodied as summation of unimodal responses, can provide a basis for optimal sensory integration [82].

However, studies of individual cell responses in multimodal areas of the cortex have revealed some important differences between sensory integration in the cortex and midbrain. Changes in bimodal relative to unimodal responses in cortex are generally smaller than those observed in midbrain, with the vast majority of changes being sub-additive, i.e. smaller than the sum of the unimodal responses [120, 136, 143-145]. Additionally, a substantial portion of individual neurons show suppression rather than enhancement of bimodal responses relative to unimodal responses. How this mixture of sub-additive enhancement and suppression of individual cell responses might contribute to the integration of sensory information remains to be shown. There is evidence that in the medial superior temporal cortex (MST) bimodal neural responses can be thought of as

sub-additive linear combinations of unimodal responses [144] and that enhancement and suppression of individual neuron responses is related to the congruence of the cells unimodal spatial tuning [136]. Gu et al. further found that neurons with congruent tuning, and hence enhancement of bimodal responses, tended to be more predictive of behavior than cells with non-congruent unimodal tuning. These results suggest that the spatial properties of responses to unimodal stimuli may be important in understanding how changes in individual neuron responses contribute to sensory integration. However, it remains to be seen whether similar explanations of mixed enhancement and suppression will hold in other multimodal areas.

While studies of sensory integration in cortex have begun to uncover distinct mechanisms of sensory integration, so far only a very limited subset of multisensory areas have been studied, and the form of sensory integration may vary across cortical areas. Many areas in the posterior parietal cortex (PPC) receive multimodal sensory inputs and are important for integrating and representing sensory information for movement [3, 5, 6, 21, 112, 114, 138, 146]. Further, studies in the superior parietal lobule (SPL) have found interactions in individual neuron responses to visual, proprioceptive and motor inputs [35, 90, 92, 126, 147, 148] suggesting that areas in the SPL play an important role in integrating these inputs.

The purpose of this study was to elucidate the neural mechanisms of integration of visual and proprioceptive reach target information in the PPC. We recorded from Area 5 and MIP, as well as nearby Area 7, in order compare integrative responses in areas that

receive different amounts of visual and proprioceptive input [6, 21, 103, 149]. We examined both individual neuron response properties in bimodal versus unimodal target reaches and the relationship of individual neuron responses to the population response in these areas. We found that individual neuron bimodal responses show a mixture of enhancement and suppression, which is unrelated to the spatial properties of the unimodal responses. However, the specific form of individual neurons' bimodal responses determine whether the population response in a given behavioral epoch and cortical area is indicative of sensory integration, i.e. gives a stronger position signal, or is similar to or smaller than unimodal population responses.

Methods

Experimental setup

Two adult male rhesus macaque monkeys (12-15kg) were used in this experiment. All procedures were approved by the UCSF Institutional Animal Care and Use Committee and followed the NIH guidelines for care and treatment of laboratory animals. As the data presented here was the basis for a previous paper, the experimental procedures for this paper have been described previously (Chapter 2). Here we summarize these procedures briefly.

The monkeys were trained to make reaches in a virtual reality setup allowing control of visual information during the task. The monkeys were head-fixed facing a mirror in which visual targets and feedback about hand position were presented. A digital video projector was used to project images onto a rear projection screen positioned above the mirror so that visual stimuli appeared in the plane of the horizontal table on which the reaching hand, the contralateral hand, rested. The target hand, the ipsilateral hand, rested on a second table positioned directly below the top table. A motor driven sleigh was used to position the ipsilateral arm at the target locations. Eye and hand position were monitored throughout the experiment. Behavioral and neural event times were recorded with a signal acquisition system that includes a programmable processor (Tucker Davis Technologies, Alachua, FL). Experiments were controlled with custom routines in Matlab (Natick, MA).

Target Modalities and Array

The monkeys were trained to reach to three different target modalities. They made reaches to visual (VIS), proprioceptive (the ipsilateral hand, PROP), and visual and proprioceptive (VIS+PROP) targets. Visual targets were filled disks 2 cm in diameter, green for VIS and blue for VIS+PROP. Proprioceptive targets were the last joint of the two middle fingers of the ipsilateral hand, moved into position with the sleigh. Reaches to these three target types were preformed for the same set of reach conditions, i.e. reach target location, fixation point, and start location.

Reach targets for all modalities were located in an arc equidistant from the projection of the monkeys' cyclopean eye position onto the tables. The exact length of the arc was determined by the extension of the ipsilateral arm (average radius 26 cm monkey C, 22 cm monkey E). Targets were positioned at 10 degree intervals with respect to the origin of the target arc (approximately 10 degrees visual angle) from -30 to +30 degrees from midline.

There were two possible fixation points, indicated with a filled red 8mm diameter disk, located at ±10 degrees from straight ahead. For each fixation point reaches were made to only six of the seven potential target locations (-30,-20,-10,0,+10,+20 for -10 fixation, -20,-10,0,+10,+20,+30 for +10 degree fixation).

All reaches were made from a visual start location, a green disk 2.4 cm in diameter, with initial feedback about the reaching hand, a white disk 1 cm in diameter at the last joint of the two middle fingers. The start target was located on the midline (measured from target arc origin: 15 cm forward monkey C, 11 cm forward monkey E).

Trial Presentation Order

The three modalities were presented in blocks so that all combinations of reach conditions were completed for PROP reaches, then VIS+PROP reaches, and finally VIS reaches before the next block (repetition) started. The modalities were separate within the blocks but the trial conditions were randomized within each modality subsection of

the block. Error trials were repeated at the end of the modality subsection within the block.

Trial and Reward Structure

In order to successfully complete a trial the monkeys had move their contralateral hand to the reach target without failing to complete any of the sequence of positional holds and delay periods enumerated here. *1) Start target acquisition*. *2) Target hand positioning:* For PROP and VIS+PROP only. *3) Fixation point acquisition*. *4) Visual target onset:* After a 700 ms fixed delay, for VIS and VIS+PROP only. Same delay for PROP. *5) Instructed delay:* Additional variable delay of 500-1000. This was the Delay behavioral epoch. *6) Go-signal:* A go-tone sounded and the start target was extinguished. *7) Reaction time:* When the hand moved 1 cm from the initial position, feedback of the hand was extinguished. *8) Movement:* The monkey had to reach without stopping to a point within a set distance from the center of the reach target (monkey C: 4 cm VIS and VIS+PROP, 5 cm PROP; monkey E: 3 cm VIS and VIS+PROP, 4.5 cm PROP). This was the Move behavioral epoch. *9) Target hold:* 200 ms. *10) Reach feedback:* On successful trials the fixation point was extinguished and visual feedback for the reaching hand was turned back on for 500 ms. On unsuccessful trials a 1 second error signal indicated which of the trial holds the monkey had violated. *11) Reward:* Monkey C received a water or fruit juice reward. Monkey E received a food reward in the form of a slurry of monkey biscuits, apple juice, and banana. Unsuccessful trials had no reward and a 1-5 second timeout before the next trial began.

To encourage accuracy, the amount of the reward depended in part on the proximity of the reach endpoint to the center of the reach target. Reaches within the inner third of the reach window received a full reward, while reaches outside this range received a reward that was half of the total reward scaled by the distance from the target divided by the size of the reach window. This reward never dropped below the minimum reward time of 50 ms (monkey C liquid reward delivered at ~1-1.2 ml/sec, monkey E slurry reward delivered at ~2.6-3.4 ml/sec). The maximum reward time was fixed within a block but increased at predetermined intervals to keep the monkeys motivated. This reward scheme resulted in very similar rewards across target types (Chapter 2).

A small portion of trials (10% in monkey C and 5% in monkey E) served as catch trials. In these trials the monkeys were rewarded without making a reach if they successfully maintained fixation through the final delay period (5), which was always extended to 1.5 seconds

Recording Cylinders

Both monkeys were trained extensively on the task before physiological recordings began. Immediately before the start of recording an 18 mm inner diameter titanium recoding cylinder, was positioned over a craniotomy opened above the IPS with the guidance of structural Magnetic Resonance Images (MRIs, monkey C: 11 mm left, -4

mm posterior; monkey E: 12 mm right, -8 mm posterior, stereotactic coordinates relative to earbars). All surgical procedures and post operative care followed NIH guidelines.

Neural Recording

Single electrode recordings were used for all data collection. All well isolated neurons that appeared modulated by the task were recorded without pre-selection for direction tuning. All neurons for which at least four blocks of trials were completed were included in further analysis. Included neurons had anywhere from 4-13 blocks with a mean of 6.2 blocks and a median of 6 blocks.

After recording, spikes were sorted and individual neurons were identified using Plexon Offline Sorter. Neurons were identified as having been recorded in Area 5, MIP or Area 7 based on the xy position of recording within the chamber and the depth of recording. The MRIs used for cylinder placement were used to make a map of the cylinder that, along with the observed neural responses to eye and hand movements, was used to identify the region from which neurons were recorded (Chapter 2). Neurons recorded less than 2000 mm from the surface of the cortex were categorized as Area 5 if they were recorded in the SPL and as Area 7 if they were recorded in the IPL. Neurons recorded below this depth in the SPL were categorized as MIP neurons. This depth cutoff was chosen because it typically corresponded to a quiet period, after the cells on the surface of the cortex, in which the electrode was likely passing through white matter.

Comparison of individual cells responses across target types

In the behavioral epochs of interest, Delay and Move, cells were tested for significant reach endpoint tuning for each modality. To do this, the firing rates were fit to a quadratic tuning curve (Equation 1) in which the rate (R) was a function of reach endpoint (T) and eye position (E) in degrees. The fit shift term (δ) allowed for eye position dependent tuning and α_1, α_2, α_3 were fit to the individual cells response. An additional term (1-γ E) was added to allow for gain modulation of responses by eye position (Equation 2) when it significantly improved the model fit across the majority of target modalities in that epoch.

$$R = \alpha_1 + \alpha_2 (T - \delta \cdot E) + \alpha_3 (T - \delta \cdot E)^2 \quad (1)$$

$$R = \left[\alpha_1 + \alpha_2 (T - \delta \cdot E) + \alpha_3 (T - \delta \cdot E)^2\right](1 - \gamma \cdot E) \quad (2)$$

Permutation tests [93] with 1000 repetitions were used to determine whether gain modulation significantly improved the tuning curve fit and whether reach endpoint tuning was significant. Inclusion of the gain term was determined by permuting the values of E used to fit gain (1-γ E) in Equation 2 across trials. Significance was determined by comparing the fraction of times the permuted SSE was smaller than the SSE with gain unpermuted. Similarly, significant endpoint tuning was determined by permuting the reach endpoints (T) across trials. The p value was computed as the fraction of times the permuted SSE was smaller than the SSE with reach endpoint information unpermuted. The epoch and modality response was significant when the p value was less than 0.05/9 =

0.0056 (Bonferroni correction, 3 target types x 3 epochs = 9). Cells were only included in further analysis if they had significant tuning in at least one modality in the epoch of interest.

Individual cell response differences across target modalities

In each epoch the two reach conditions with the highest firing rates (the peak of tuning) and the two reach conditions with the lowest firing rates (the minimum of tuning) were selected using the average firing rates across modalities. We have previously shown that tuning curves are generally highly correlated across the three target modalities (Chapter 2). Thus, selecting peak and minimum conditions from the average rate provided a good measure of changes in the response across target modalities and gave qualitatively the same results as comparisons made with peak and minimum selected separately across modalities. Significant differences across modalities at the peak and minimum as well as significant differences in modulation (peak - minimum) were determined using a permutation test with 5000 repetitions.

Individual cell firing rate decode

In order to obtain a measure of how predictive a cells' firing rate in a given modality and epoch (modality-epoch-cell) was of behavior we used a linear decode of firing rate. To do this we used a linear regression where firing rate was the independent variable and reach endpoint, with the dependence on eye position (δ) determined from

Equation 1, was the dependent variable ($T - \delta E$). We used leave-one-out cross validation across trials to compute an R^2 value for this regression. The cross validation R^2 value then served as a measure of how predictive firing rate was of reach endpoint.

Comparison of population responses across target modalities

In order to look at how specific types of changes in the tuning curve for bimodal versus unimodal responses affected population coding, we rearranged each cells responses so that all cells could be thought of as having the same canonical tuning curve. To do this we ranked each cells' responses for the twelve task conditions (6 targets for each of 2 fixation points) from lowest to highest firing rate, based on the average response across modalities (the results of this analysis were unchanged if the responses were ranked separately for each target modality). The responses to for each of the twelve conditions were then reordered to follow this ranking for the cell. This aligned all cells' firing rates in order from minimum response to maximum response and allowed us to compare how changes in each part of the tuning curve affected the population response. The assumption here is that if we had measured all of the neurons in the area the population response would be made up of many groups of cells with similar tuning. These groups of neurons would form the population code for target and movement information in these areas. By rearranging our cells responses we are pretending that we have sampled a single group with the same tuning rather than a subset of cells with many different preferred directions. The changes in the population response for this resampled

group, which now shares a single preferred direction, should in theory occur in all groups coding the same preferred direction.

In order to compare changes across the population each cells' responses (irrespective of modality) were normalized so the responses across Delay, Reaction time, Move fell between zero and one. The population tuning for the VIS+PROP targets was then compared either to the highest unimodal task response (VIS or PROP) chosen independently for each cell, or to the population responses during the VIS and PROP tasks. In each case the ranked and aligned responses were fit with linear regression to a line. A paired permutation test with 5000 repetitions was used to determine whether the fit parameters were significantly different between the population responses being compared.

Results

Behavior across target types

The behavioral performance was comparable across target types, and has been reported in detail previously (Chapter 2). Briefly, both monkeys preformed smooth quick reaches to all three target types with the distributions of reaction times and peak velocities overlapping across target types. Additionally, the graded reward scheme (see Methods) successfully minimized differences in the rewards across target types. The

average rewards differed by less than 30 ms (~3 μl at ~1 ml/sec monkey C, ~10 μl at ~3 ml/sec monkey E) across target types for both animals, while the standard deviation in reward size was more than 130 ms for all target types. Thus, while there are small behavioral and reward differences across target types, these differences are much smaller than the natural variation within a target type, making them unlikely to play a large role in the neural responses reported here.

One important, and comparatively large, difference between the three target types was reach endpoint variance. Visual and proprioceptive information have different amounts of intrinsic variability [60, 70]. Providing target information with vision or proprioception should lead to changes in reach endpoint variance that reflect both the computations needed to plan the reach and the reliability of the sensory modality [70, 107]. This change in endpoint variance across target types is readily apparent in both monkeys' data, with VIS variance being considerably smaller than PROP variance (Figure 1). If the monkeys are integrating sensory information, when both visual and proprioceptive information are available, the reach endpoint variance should be smaller than the unimodal variances [57, 63, 70]. Comparing reach endpoint variance across days of recording reveals that this is the case; the VIS+PROP variability is generally smaller than both the VIS (Figure 1A) and the PROP (Figure 1B) variability. Additionally, changes in the day-to-day VIS+PROP variance are correlated with both the VIS and PROP variance (see Figure 1A&B for R values). These correlations are higher than the day-to-day correlation between VIS and PROP variance ($R=0.195$ $p=0.043$ monkey C, $R=0.386$ $p=0.001$ monkey E), indicating that the monkeys are using both sources of sensory information to

plan and execute their reaches. Further, while the decrease in VIS+PROP variance relative to VIS variance was relatively small, it was highly significant ($p<<0.0001$, t-test) and followed a similar pattern to the decrease in variability seen in humans performing the same tasks (compare Figure 1C&D). Thus, the changes in reach endpoint variance seen across the three target types support the conclusion that the monkeys are using and integrating the sensory information available in a statistically principled manner.

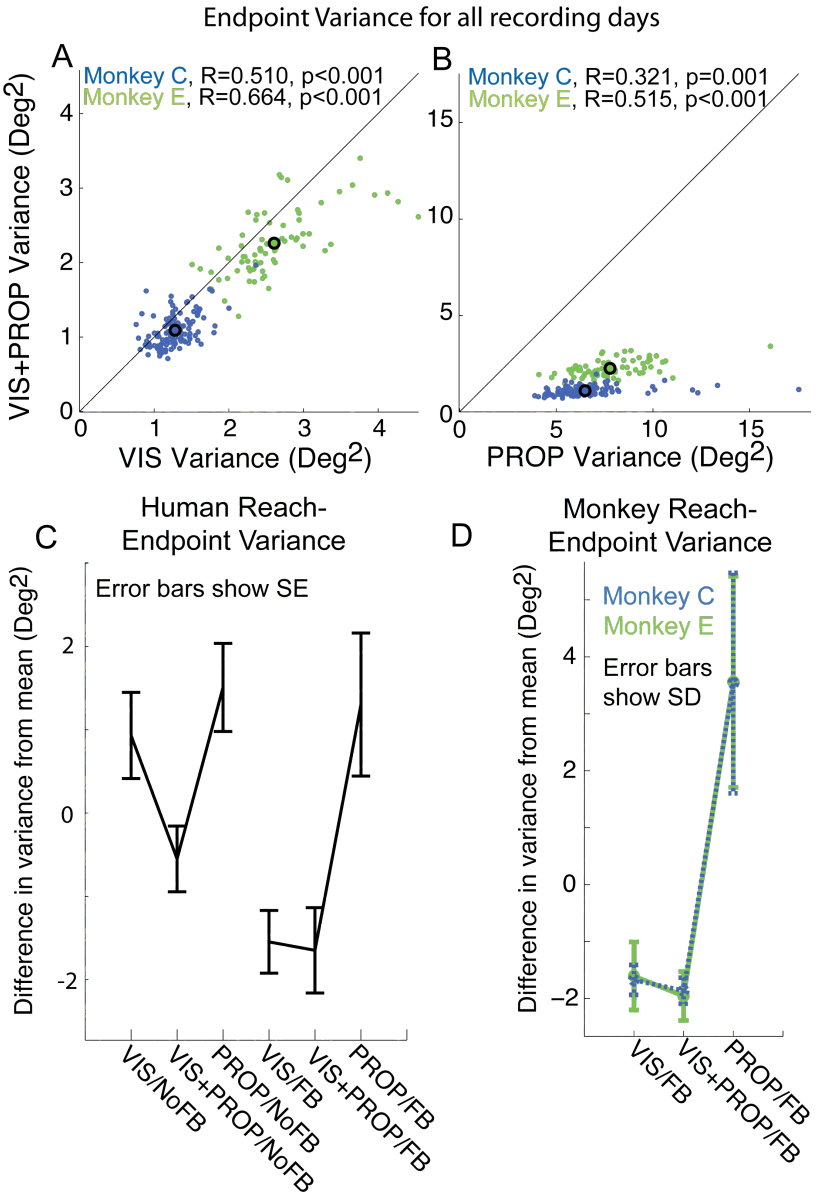


Figure 1: Reach endpoint variance across target modalities. A) Comparison of reach endpoint variance in VIS and VIS+PROP for all recording days. B) Comparison of reach endpoint variance in PROP and VIS+PROP for all recording days. Each point is the average variance on a single day of recording. Large circles show average variance across days. C) Average changes in reach endpoint variance for humans performing reaches to VIS, VIS+PROP, and PROP targets with (FB) and without (NoFB) initial visual feedback of the reaching hand (from [107]). D) Changes in reach endpoint variance for monkeys C and E during VIS, VIS+PROP, and PROP FB reaches. VIS+PROP variance is significantly smaller than PROP or VIS variance ($p<0.0001$, t-test).

Neuron responses in the PPC

We recorded a total of 446 neurons from the PPC for this experiment. Of these cells 193 were nominally in Area 5 (101 monkey C, 92 monkey E), 182 were nominally in MIP (95 monkey C, 87 monkey E), and 71 were nominally in Area 7 (monkey E). We will focus primarily on the Area 5 and MIP responses. A complete breakdown of the tuning properties across target type and epoch of the cells recorded from Area 5 and MIP, as well as the presumed recording locations for these cells, has been reported previously (Chapter 2). Cells were included in the analyses presented here if they had significant reach endpoint tuning for at least one target type in the epoch of interest (see Methods).

We first asked how neural activity differed for the two unimodal target types, VIS and PROP, in Area 5 and MIP. We have previously shown tuning properties are similar across VIS and PROP reaches for individual cells (Chapter 2). Thus, we quantified differences between VIS and PROP by comparing the modulation, the difference between the peak and minimum firing rate across all reach conditions, across target types. The majority of cells in Area 5 and MIP did not have significant differences in modulation ($p<0.05$) between VIS and PROP reaches (68% Area 5 Delay, 58% MIP Delay, 72% Area 5 Move, 55% MIP Move). The cells that did have significant differences between VIS and PROP modulation showed heterogeneous modality preferences, though there were significant differences in average modulation across tasks. During Delay Area 5 cells showed significantly greater modulation during PROP reaches across the population (Figure 2A, $p=0.003<0.05/2=0.025$, permutation test, Bonferroni correction, 2 epochs),

while MIP cells showed significantly greater modulation during VIS reaches ((Figure 2B, $p<0.001<0.025$, permutation test, Bonferroni correction). During Move the Area 5 population lost its tendency to be more modulated by PROP reaches, showing if anything a slight preference for VIS reaches, though this tendency was not significant (Figure 2C, $p=0.053$, permutation test). The MIP population maintained its preference for VIS reaches during Move (Figure 2D, $p<0.001<0.025$, permutation test, Bonferroni correction). Thus, we found that both Area 5 and MIP have strong responses to VIS and PROP targets, though Area 5 shows a slight tendency to respond more strongly to PROP, and MIP shows a tendency to respond more strongly to VIS, which is consistent with the different strengths of the visual and somatosensory inputs to these areas [3, 103, 115].

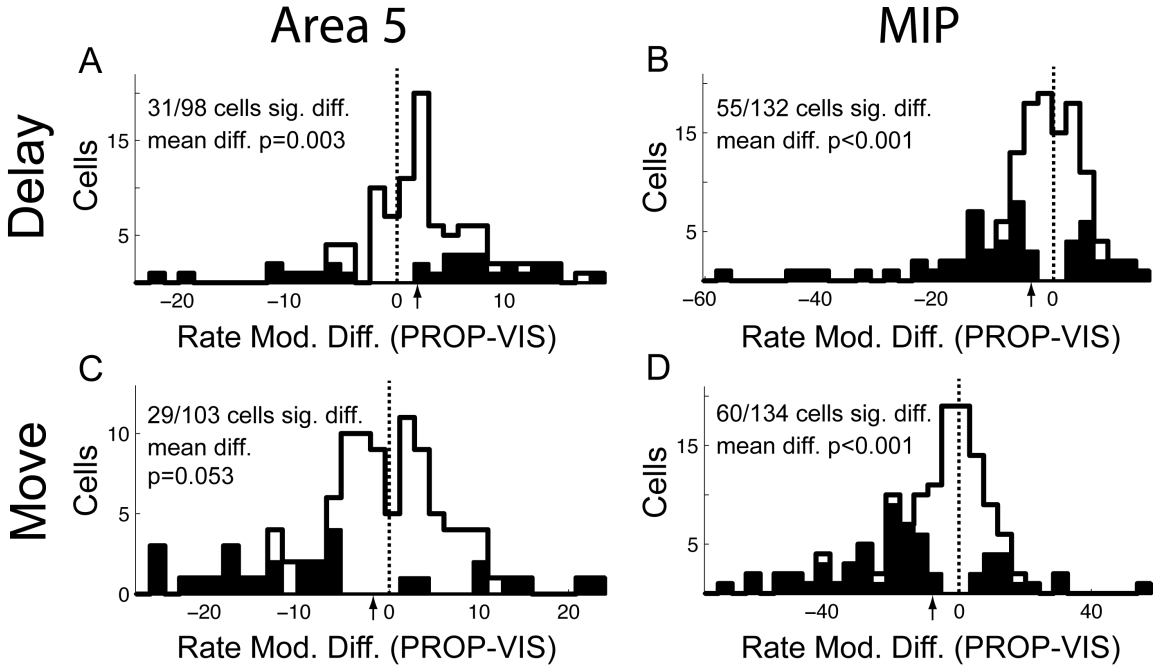


Figure 2: Histograms of differences in firing rate modulation in VIS versus PROP reaches in Area 5 and MIP for Delay and Move. Filled histograms show cells with significant differences ($p<0.05$, paired permutation test) in modulation between target modalities. Arrows indicate average difference. P-values for mean differences come from paired permutation tests.

Modulation of single cell responses by bimodal target information

We first looked for changes in individual neuron's instructed delay period (Delay) responses when bimodal target (VIS+PROP) responses were compared to unimodal target (VIS or PROP) responses. As Delay period responses in the PPC are generally thought to correspond to movement planning [4, 6, 38, 117, 137] neural correlates of the behavioral decrease in variance for bimodal targets seen in our task (Figure 1) should be apparent in these responses.

There were a variety of changes in individual cells Delay VIS+PROP responses relative to their unimodal responses in both Area 5 and MIP (Figure 3 shows several example cells). Bimodal sensory integration on the single cell level is generally defined as a difference between the cells bimodal response and its best (VIS or PROP, higher firing rate at peak) unimodal response [150]. Some cells displayed enhancement at the peak of the tuning curve (Figure 3A, $p=0.014$, permutation test), which could provide a more reliable position signal during the VIS+PROP reaches. However, cells also showed significant suppression of VIS+PROP responses relative to the best unimodal response. This suppression could be general suppression relative to the best unimodal response (Figure 3B $p<0.001$ at peak and $p=0.003$ at minimum, permutation test) or specific to either the peak of the tuning curve (Figure 3C, $p<0.001$, permutation test) or, less frequently, the minimum of the tuning curve. Additionally, the majority of cells had similar responses to the best unimodal and VIS+PROP targets (Figure 3D, $p=0.299$ at peak and $p=0.292$ at minimum, permutation test). These responses were matched to the

VIS response in some cases and the PROP response in others. Thus, individual cells showed enhancement, suppression, or no change in response for bimodal versus unimodal targets.

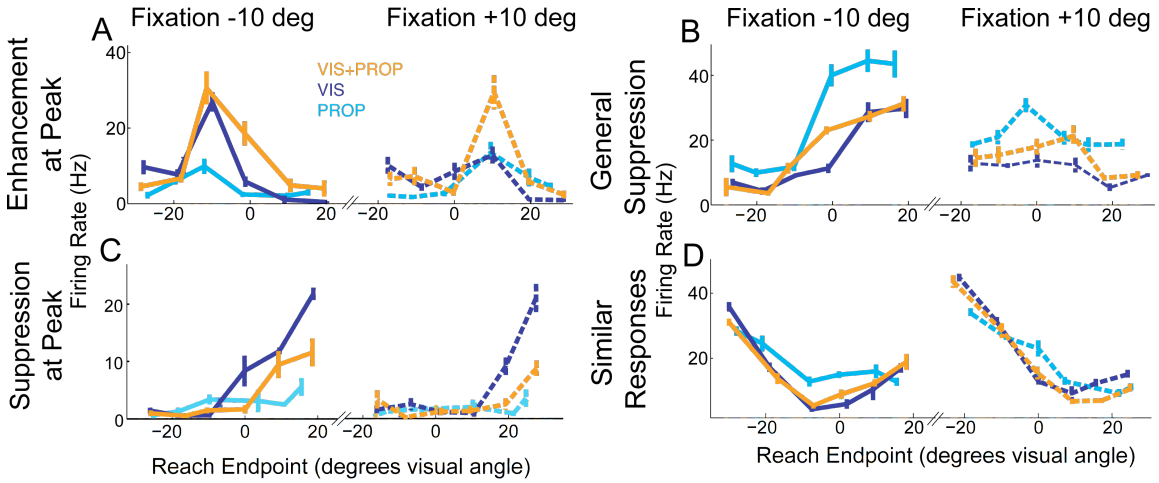


Figure 3: Example cells' responses across the three target modalities. Responses for the two fixation points are separated for clarity. A) Cell shows enhancement of VIS+PROP responses at the peak of the tuning curve ($p=0.014$). B) Cell shows suppression of VIS+PROP responses relative to the best unimodal (PROP) responses at the peak ($p<0.001$) and minimum ($p=0.003$) of the tuning curve. C) Cell shows suppression of VIS+PROP responses relative to the best unimodal (VIS) responses at the peak of the tuning curve ($p<0.001$). D) Cell has similar responses for VIS+PROP and best unimodal (VIS) reaches ($p=0.299$ at peak, $p=0.292$ at minimum. P-values determined with paired permutation tests (see Methods). Error bars show standard error.

We next asked whether there were any specific types of changes in response that were more common than others across the population. As changes in the peak and the minimum of tuning provide a proxy for measuring changes in the neurons responsiveness, and many neurons seemed to have changes specific to these parts of the tuning curve, we tested all cells for significant differences in peak or minimum between VIS+PROP and the best unimodal response ($p<0.05$, 2 comparisons no correction for multiple comparisons see methods). Of the cells tuned for Delay, 28/98 Area 5 cells and 52/132 MIP cells showed either significant enhancement or suppression at either the peak

or the minimum of the tuning curve. We also examined responses during the movement epoch (Move), and found that 42/103 Area 5 cells and 56/136 MIP cells showed either significant enhancement or suppression at either the peak or the minimum of the tuning curve. In general all changes in modulation were subadditive, i.e. less than the sum of the unimodal responses (Supplemental Figure 1). While both enhancement and suppression were common, there were subtle differences in the distribution of these effects between the peak and minimum of the tuning curve in both areas during Delay and Move (Supplemental Figure 1). Thus, we next set out to determine how changes in different parts of individual cells' tuning curves will affect the population response for bimodal versus unimodal targets.

Population Delay responses to VIS+PROP reaches in Area 5 and MIP

Changes in response at different parts of the tuning curve will have different implications for information coding within the cell. For example, selective suppression at the peak of the tuning curve would in theory provide less information, while selective suppression at the minimum of the tuning curve would in theory provide more information. This is because these two changes would have opposite affects on the cells overall modulation. In order to look for changes in the VIS+PROP response that were localized to different parts of the tuning curve, we aligned all of the cells responses so that we could compare changes across the tuning curve. In order to do this we ranked the firing rates of each cell on all reach conditions from the lowest response the highest response (1-12, for each of the 6 target x 2 fixation conditions) and aligned the conditions

in order of their mean rate (see Methods). This allowed us to treat the entire population of cells as though they had the same tuning and look for changes in specific parts of the population tuning curve in response to bimodal sensory information. If we could sample the entire population of cells within an area, the population response would be made up of many groups of cells with similar preferred directions and tuning curves. By aligning all of our cells so that they have the same nominal tuning curve we can use our sample as a proxy for changes in population activity across all preferred tuning directions.

We first looked for consistent changes in individual cells' Delay responses to bimodal targets relative to their best unimodal response. Unimodal task preferences were heterogeneous in both Area 5 and MIP (Figure 2), so the best unimodal response was chosen separately for each cell as described above. When the average VIS+PROP response was compared to the average across all cells' best unimodal responses, a slight but significant trend towards suppression of bimodal responses at the minimum of the tuning curve was seen in both Area 5 and MIP (gold versus grey lines, Figure 4A&C, $p<0.001<0.05/8=0.006$, paired permutation test, Bonferroni correction, 2 line parameters x 2 comparisons x 2 sets of comparisons). In contrast, changes in individual cells' responses at the peak of the tuning curve had no net effect on VIS+PROP responses relative to the average of cells' best unimodal responses ($p>0.5$, paired permutation test). Thus, in both Area 5 and MIP many of the changes in individual neurons bimodal responses relative to their unimodal responses have no net effect on the population response. There are, however, small localized changes in tuning curves across the population, which could increase the amount of information present for bimodal stimuli.

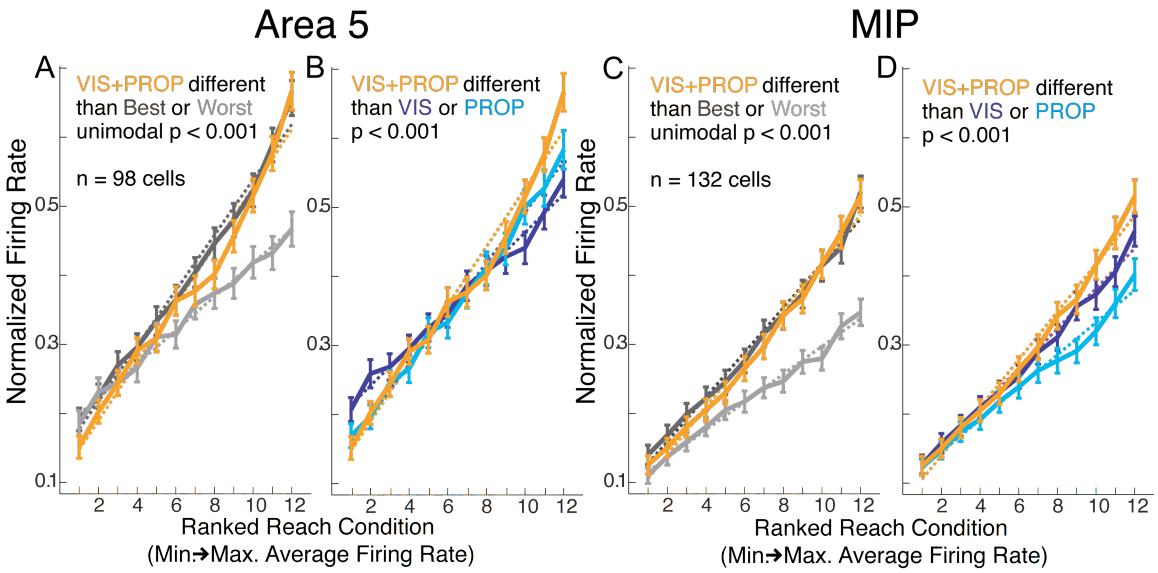


Figure 4: Delay population responses in Area 5 and MIP. Responses for each cell are ordered from lowest firing rate reach condition to highest firing rate condition. Each cells' firing rate is normalized so that it ranges from zero to one in its Delay, React and Move responses (the same normalization is used across target modalities). A&C) Show average VIS+PROP responses compared to the average of the best and worst unimodal responses chosen for each cell. B&D) Show average VIS+PROP responses compared to average VIS and PROP responses. Error bars show standard error. Dashed lines show linear fit to average responses. P-values are from paired permutation tests for significantly different fits across average responses.

However, comparing the population responses across target types provides a much clearer picture of an area's population than comparing individual cells' bimodal and best unimodal responses. Individual cells' responses to VIS and PROP targets were often very similar (Figure 2) and the choice of best unimodal response in these cases was within measurement noise. Further, as both Area 5 and MIP had heterogeneous unimodal preferences when there were significant differences across modalities, this is illustrated by the fact that the average Delay responses for VIS and PROP fell in between the average of the best and worst unimodal responses (compare Figure 4 A&C to B&D, though Area 5 was on average more modulated for PROP, and MIP was on average more modulated for VIS, $p<0.001$, paired permutation test). This distinction between mean VIS

and PROP and average of the best unimodal responses alters the picture of bimodal responses across the population. The mean VIS+PROP responses were significantly greater than VIS or PROP mean responses near the peak of the tuning curve in both Area 5 and MIP (Figure 4B&D $p<0.001<0.008$, permutation test, Bonferroni correction). In contrast, the mean VIS+PROP and unimodal responses are similar at the minimum of the tuning curve (similar to PROP in Area 5, similar to VIS and PROP in MIP, $p>0.6$, paired permutation test). Thus, at the peak, where we saw no change in mean VIS+PROP relative to cells' best unimodal response, we see enhancement relative to VIS and PROP. At the minimum, where we saw selective suppression relative to best unimodal response, we see no change relative to unimodal. This suggests an interaction between changes in individual cells bimodal relative to unimodal responses and the heterogeneous task preferences in Area 5 and MIP. This interaction results in increased VIS+PROP modulation across the population in both Area 5 and MIP. This could provide a stronger position signal for downstream areas to account for the decrease in reach endpoint variance seen in our behavioral data (Figure 1).

The increase in mean VIS+PROP response does not depend on reordering responses or on having cells that respond to only one modality. When cells' responses are not reordered to align their tuning curves, the average VIS+PROP responses are similar to or greater than VIS or PROP responses for all reach conditions (Supplemental Figure 2A&B). Additionally, when small groups of cells with the same preferred direction are averaged, the same trend towards increased VIS+PROP response at the peak of the tuning curve can be seen (Supplemental Figure 2C-F). This suggests that our reordering results

are indeed representative of the population response for a large group of cells with the same tuning. Further, the increased VIS+PROP modulation seen in Figure 4 is present even when only cells tuned for all three target types are included in the population average (data not show). Thus, this effect is not the trivial result of separate VIS and PROP populations being averaged together. All of these findings support the idea that the difference between mean VIS+PROP and mean unimodal responses is a specific increase in bimodal modulation across a multimodal population of neurons.

It should be noted that either the selective suppression of the minimum (Figure 4A&C) or the selective enhancement of the peak (Figure 4B&D) would increase modulation of cells' responses. To examine the role of changes in individual cells' bimodal responses relative to the role of heterogeneous target type preferences we next examined the modulation of individual cells in Area 5 and MIP. We compared the VIS+PROP Delay modulation to the modulation in the best unimodal target response, and to either VIS modulation or PROP modulation (Figure 5). While the VIS+PROP responses show a slight trend towards increased modulation relative to the best unimodal response for each cell, this trend fails to reach significance in either area (Figure 5A&B, $p=0.052$ in Area 5, $p=0.097$ in MIP, paired permutation test). In contrast the VIS+PROP responses showed on average greater modulation than either the VIS (Figure 5C&D, $p<0.001<0.05/6=0.008$, paired permutation test, Bonferroni correction) or the PROP (Figure 5E&F, $p<0.001<0.008$, paired permutation test, Bonferroni correction) responses in both Area 5 and MIP. This suggests that the heterogeneous modality preferences in Area 5 and MIP play a larger role in the increased VIS+PROP modulation than changes

in the individual cell responses, though both may be important for the integration of sensory information.

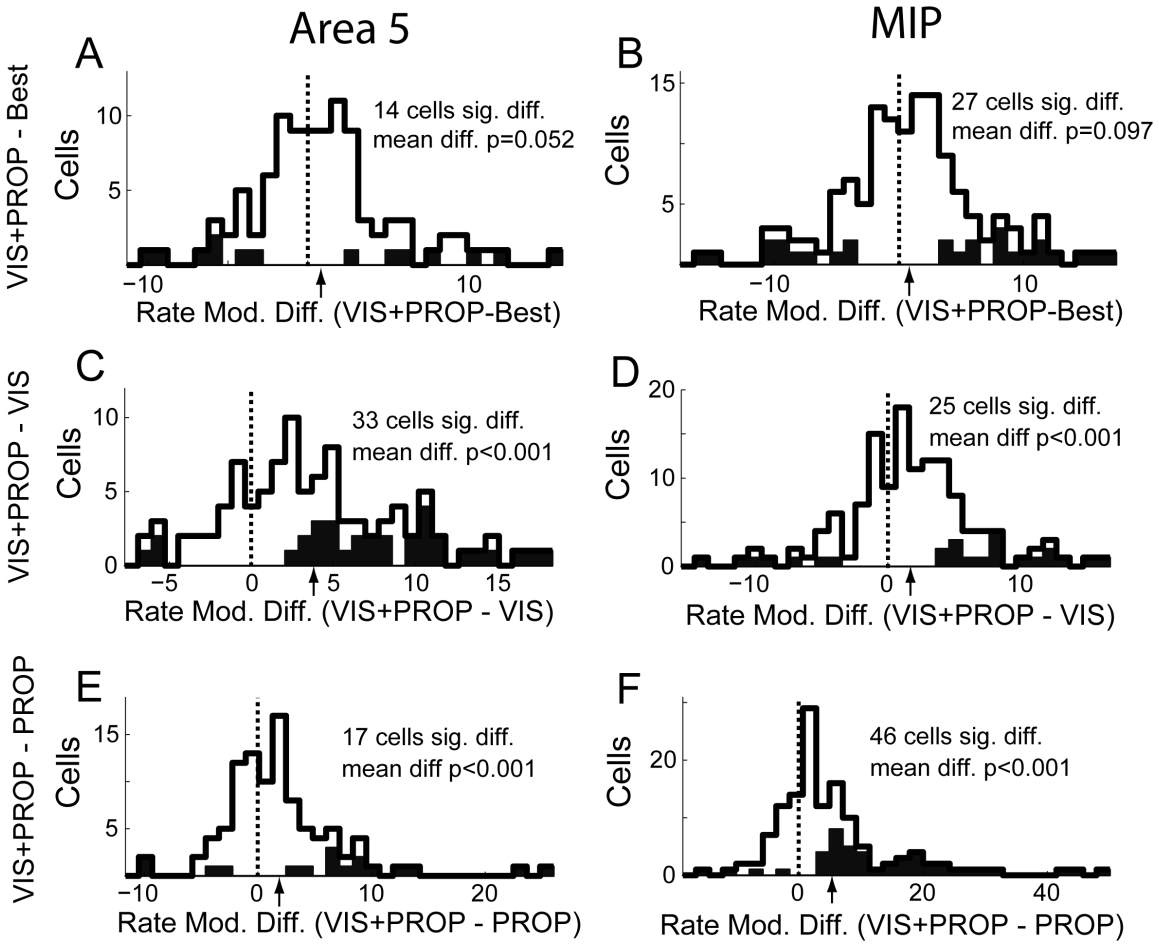


Figure 5: Histograms of individual cell differences in Delay modulation for Area 5 and MIP. A&B) Differences between VIS+PROP and each cells best unimodal response modulation. C&D) Differences between cells VIS+PROP and VIS modulation. E&F) Differences between VIS+PROP and PROP modulation. Filled histograms show individual cells with significant differences. Arrows indicate mean difference in response. P-values for mean difference come from paired permutation tests.

Correlates to changes in Delay modulation during VIS+PROP reaches

We have used changes in modulation as an indicator for how much information about target location is being encoded by both individual cells and across the population.

Here we examine more closely how changes in firing rate modulation relate to the alignment between cells unimodal tuning curves and the information that can be decoded for each target type about reach endpoint.

Studies in other multimodal areas have shown a relationship between whether individual cells' bimodal responses were enhanced or suppressed and the tuning properties of the unimodal responses [136]. Specifically, cells with similar tuning across the unimodal sensory inputs tended to show enhancement of bimodal responses, while cells that showed divergent tuning across unimodal inputs tended to show suppression of bimodal responses. While we have previously shown that tuning curves tend to be highly correlated across VIS and PROP reaches (Chapter 2) this alignment of responses is not perfect, and it was possible that slight variations in tuning across visual and proprioceptive inputs could account for the mixture of enhancement and suppression observed here. However, when the correlation between individual cells VIS and PROP tuning curves was compared to the changes in modulation in VIS+PROP relative to the best unimodal response no relationship between these factors was apparent (Figure 6A, $R= -0.025$, $p>0.790$, Spearman's correlation, combined across areas as the results were similar for Area 5 and MIP). This indicates that the mixture of enhancement and suppression of individual cell responses cannot be explained by differences in tuning across the unimodal tasks.

We next wanted to verify that the changes in modulation reflected changes in the information encoded by individual cells. In order to test this we a used a linear decoding

model of each cells' trial by trial firing rate to look at the reach endpoint information that could be read out for each target type (see Methods). We then compared changes in modulation across target types to changes in how predictive the cells' firing rate was of reach endpoint (Figure 6B). There is a significant positive correlation between changes in modulation between the VIS+PROP and best unimodal response and changes in predictiveness across the VIS+PROP and best unimodal responses (Figure 6B, R=0.382, $p<0.001<0.05/3=0.016$, Spearman's correlation, Bonferroni correction). Thus, enhancement and suppression of modulation in VIS+PROP relative to the best unimodal responses is related to the information that can be read out from the cells' firing rate. However, the fact that this relationship lead to both increases and decreases in information indicates that additional factors, such as the heterogeneous target type preferences seen in both Area 5 and MIP, must be involved to obtain the increase in target position information suggested by the behavioral decrease in endpoint variance for bimodal targets.

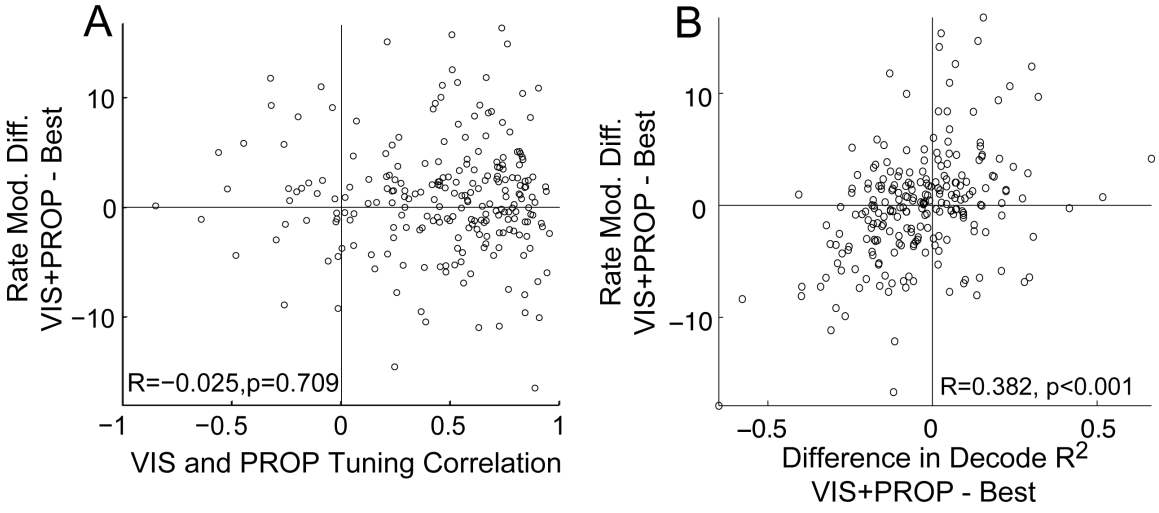


Figure 6: A) Comparison of difference between VIS+PROP and best unimodal modulation to correlation in unimodal tuning curves for individual cells Delay responses. B) Comparison of difference between VIS+PROP and best unimodal modulation to difference decode R^2 value

between VIS+PROP and best unimodal Delay responses. R and p-values are from Spearman's correlation.

Population Delay responses to VIS+PROP reaches in Area 7

There are several factors that contribute to the increase in VIS+PROP modulation seen in Area 5 and MIP and make this a feasible mechanism for sensory integration in these areas. First, changes in bimodal versus best unimodal response are small across the population and localized to selective suppression of the minimum of the tuning curve (Figure 4A&C). Second, unimodal target preferences are heterogeneous in these areas resulting in strong responses to both visual and proprioceptive target information (Figure 2 and Figure 4B&D). Third and finally, the tuning curves across all target types are roughly aligned (Chapter 2 and Figure 6A) so downstream areas can read out positional information across the whole population irrespective of the sensory modality specifying target information. Here we compare responses in Area 5 and MIP to responses in Area 7 and show that when the first two conditions are not met there is no increase in bimodal responses.

When average VIS+PROP Delay responses in Area 7 are compared to either best unimodal or VIS and PROP responses, with the same analysis used for Area 5 and MIP responses above, a different pattern of bimodal changes in response emerges. In Area 7 the mean VIS+PROP response showed general suppression relative to the mean of each cells best unimodal response (Figure 7A, $p=0.003<0.05/8=0.006$, paired permutation test, Bonferroni correction). This is in contrast to the selective suppression seen in Area 5 and

MIP (Figure 4A&C). In addition cells in Area 7 respond much more strongly for reaches to VIS than PROP targets ($p<0.001$, paired permutation test), and when the population response is compared across target types, the mean VIS+PROP response in Area 7 is significantly suppressed relative to the mean VIS response (Figure 7B, $p=0.001<0.006$, paired permutation test, Bonferroni correction). Though the suppression of VIS+PROP responses relative to VIS responses appears greater at the base of the tuning curve than at the peak (Figure 7B), there is no significant difference between modulation of VIS+PROP and VIS responses in Area 7 ($p=0.870$, paired permutation test), suggesting that the difference between VIS+PROP and VIS responses is consistent with general suppression of responses. Thus, Area 7 does not display the localized changes in bimodal versus best unimodal response or the heterogeneous task preferences seen in Area 5 and MIP. The absence of these effects results in changes in bimodal responses in Area 7 that appear more consistent with suppression due to competition between sensory inputs than integration resulting increased modulation of bimodal responses.

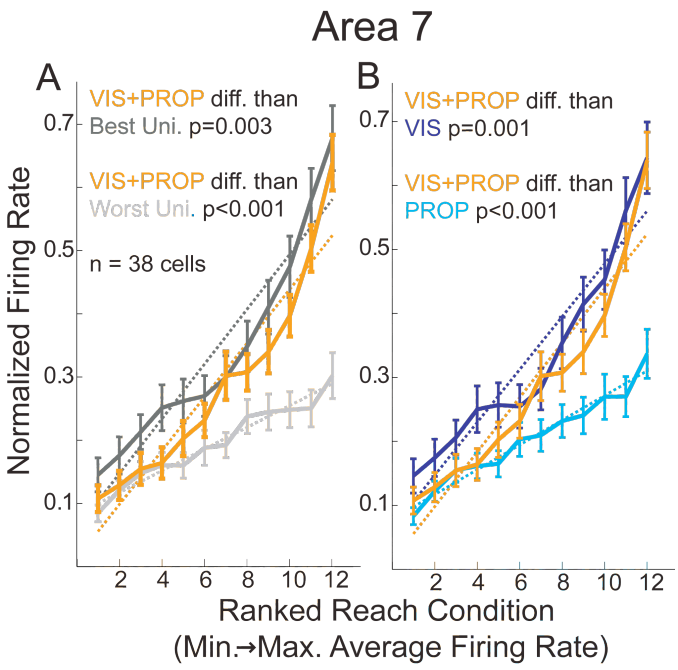


Figure 7: Delay population responses in Area 7. Responses for each cell are ordered from lowest firing rate reach condition to highest firing rate condition. Each cells' firing rate is normalized so that it ranges from zero to one in its Delay, React and Move responses (the same normalization is used across target modalities). A) Average VIS+PROP responses compared to the average of the best and worst unimodal responses chosen for each cell. B) Average VIS+PROP responses compared to average VIS and PROP responses. Error bars show standard error. Dashed lines show linear fit to average responses. P-values are from paired permutation tests for significantly different fits across average responses.

Population Move responses to VIS+PROP reaches in Area 5 and MIP

Examining population responses during Move in Area 5 and MIP reveals a divergence between planning and movement activity, and illustrates the importance of the specific type of changes in individual cells' bimodal versus best unimodal responses. During Move the average bimodal relative to best unimodal response is generally suppressed, with greater suppression at the peak of the tuning curve, in both Area 5 and MIP (Figure 8A&C, $p<0.001<0.05/8=0.006$, paired permutation test, Bonferroni correction). This is in contrast to the selective suppression at the minimum seen during Delay. This general suppression of bimodal responses, combined with these areas' heterogeneous modality preferences (Figure 2), results in similar mean responses to VIS+PROP and unimodal reaches across the population (similar to VIS and PROP in Area 5, similar to VIS in MIP, Figure 8B&D, $p>0.015>0.006$, paired permutation test, Bonferroni correction). The similar bimodal and unimodal target responses suggests that the selectivity of the suppression seen during Delay is key to the increase in bimodal modulation during movement planning, and that, unlike Delay, Move may not encode stronger positional information for bimodal targets.

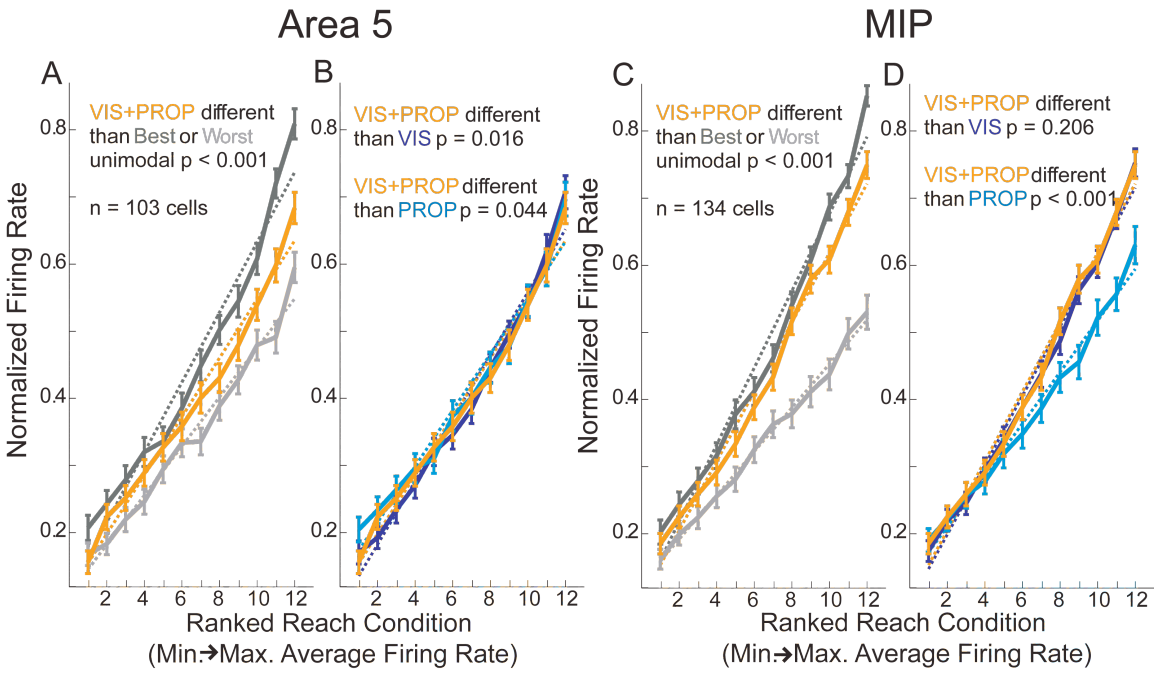


Figure 8: Move population responses in Area 5 and MIP. Responses for each cell are ordered from lowest firing rate reach condition to highest firing rate condition. Each cells' firing rate is normalized so that it ranges from zero to one in its Delay, React and Move responses (the same normalization is used across target modalities). A&C) Show average VIS+PROP responses compared to the average of the best and worst unimodal responses chosen for each cell. B&D) Show average VIS+PROP responses compared to average VIS and PROP responses. Error bars show standard error. Dashed lines show linear fit to average responses. P-values are from paired permutation tests for significantly different fits across average responses.

The changes in average Move modulation are related to changes in individual cells' modulation. Individual cells' VIS+PROP Move modulation is on average smaller than their best unimodal modulation in both Area 5 and MIP (Supplemental Figure 3A&B, $p<0.014<0.05/3=0.017$, paired permutation test, Bonferroni correction). However, there are on average no significant differences in individual cells' modulation between VIS+PROP and VIS or PROP in Area 5 (Supplemental Figure 3C&E, $p>0.13$, paired permutation test, Bonferroni correction) or between VIS+PROP and VIS in MIP (Supplemental Figure 3D, $p>0.13$, paired permutation test, though VIS+PROP responses are more modulated than PROP responses due to this areas' preferences for VIS targets,

Supplemental Figure 3F). This suggests that the general suppression of VIS+PROP versus best unimodal responses serves to equalize responses between VIS+PROP and VIS or PROP responses in Area 5, and VIS responses in MIP. The lack of changes in individual cells' bimodal versus unimodal modulation supports the idea that cells in Area 5 and MIP are not providing additional information about bimodal target location during Move.

The change from increased modulation in VIS+PROP during Delay to similar modulation in VIS+PROP and unimodal responses during Move emerges gradually over the course of the trial (Figure 9). When Delay tuned cells are aligned on the Go-Signal, the increased VIS+PROP modulation is significant before the movement starts, during the instructed delay, but decreases after the Go-Signal, as the movement approaches, in both Area 5 and MIP (Figure 9A&C). When Move tuned cells are aligned on the start of movement, VIS+PROP modulation is not significantly different than VIS or PROP modulation in Area 5 and not significantly different than VIS modulation in MIP, both during the reaction time before the movement and after the initiation of the movement (Figure 9A&C). Thus, the change in modulation of VIS+PROP target responses appears to emerge during the transition from planning to execution of the movement. This suggests that the change from movement planning to reach feedback results in differential processing of multisensory information.

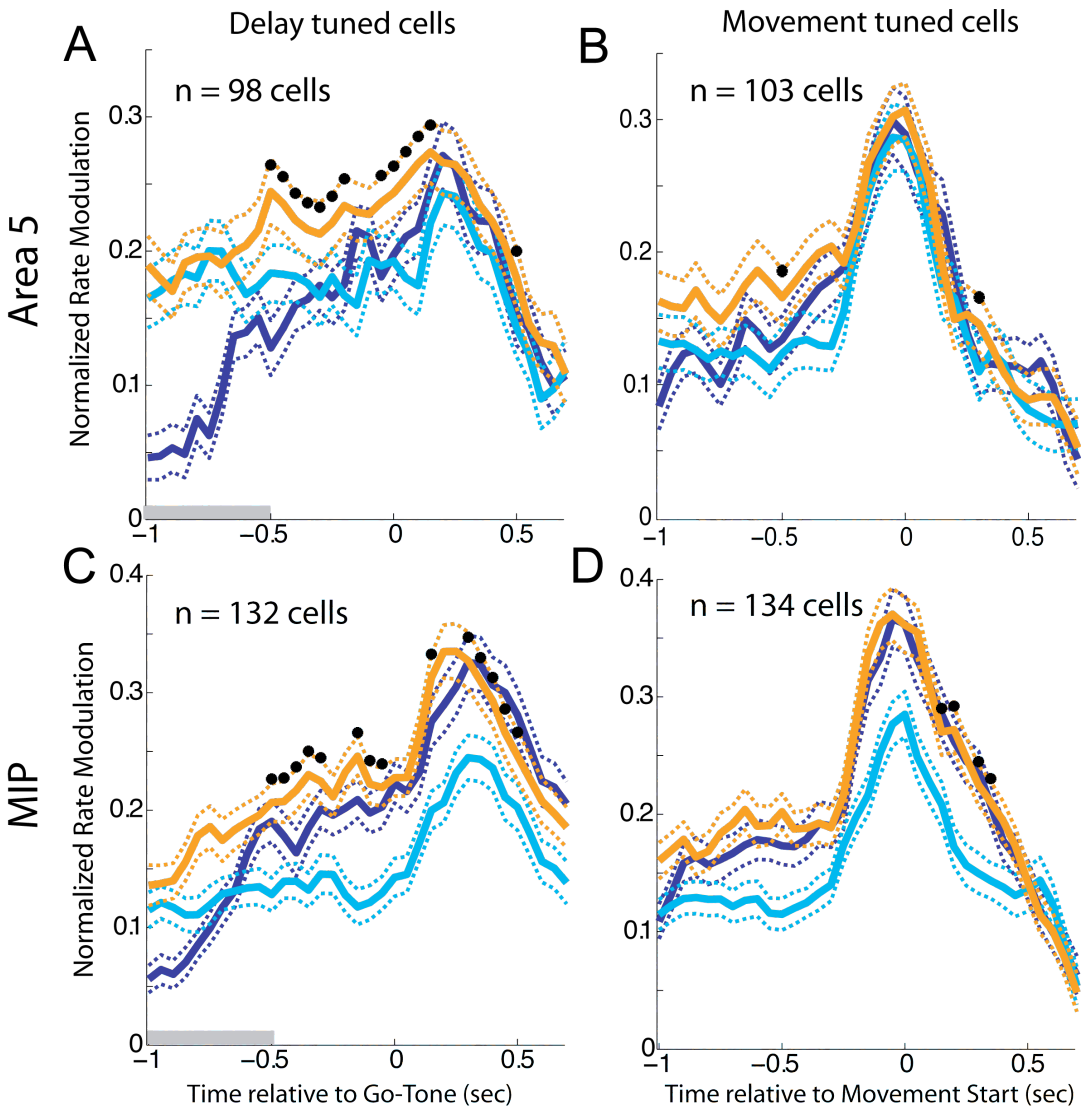


Figure 9: Peristimulus time histograms of population modulation. A) Delay tuned Area 5 cells aligned on Go-Tone. C) Delay tuned MIP cells aligned on Go-Tone. Grey bars in A&C indicate period during which the visual target appeared. B) Move tuned Area 5 cells aligned on movement start. D) Move tuned MIP cells aligned on movement start. Each cells response is normalized over the range of the PSTH. Dashed lines indicate standard error. Stars indicate time bins when the VIS+PROP modulation was significantly different than VIS and PROP modulation ($p<0.05$, paired permutation test on line fit to ranked and ordered responses). Differences in modulation were only tested for the 500 ms before and after the aligning cue. Average reaction time across both animals was 370 ms.

Discussion

This study examined how the responses of neurons in the PPC change when bimodal sensory information, rather than unimodal information, about target position is available. Specifically we looked at how visual and proprioceptive, in this case the ipsilateral hand, target information affect the responses of both individual neurons and the population during reaching. During instructed delay, we found that many neurons in Area 5 and MIP show enhancement or suppression of bimodal responses relative to their best unimodal responses. The exact form of these changes varied greatly from neuron to neuron, and only a slight suppression of responses at the minimum of the tuning curve was consistent across the population in these areas. However, this specific suppression of minimum responses, combined with the areas heterogeneous task preferences, resulted in increased bimodal modulation across the population in both Area 5 and MIP during Delay. By movement initiation, the increased bimodal modulation disappeared, suggesting that the roles of Area 5 and MIP in sensory integration for planning and feedback may be different. The interaction between changes in bimodal responses in single cells and the population across the planning and movement periods in this study suggest a novel way of thinking about sensory integration within a cortical area.

One factor that likely played a role in the types of responses seen in this study was the decision to use the ipsilateral and not the contralateral hand as the proprioceptive target. This decision was motivated by the desire to study the way in which visual and proprioceptive information are used, both separately and jointly, to guild reaching. Neural

responses in the contralateral hemisphere of the PPC have been shown to be more correlated with reach reaction time than responses in the ipsilateral hemisphere [128], suggesting that the contralateral responses are more closely linked to movement. However, areas within the PPC respond to ipsilateral stimuli and hand movements [111-114, 127, 128], and substantial bimanual responses to somatosensory stimuli have been reported in somatosensory cortex adjacent to Area 5 [151, 152]. Further, both the premotor and motor cortices have neurons that respond to ipsilateral and bilateral hand movements [153-157]. Thus, coordinating movements between the two hands involves both hemispheres of many motor and sensory-motor areas, which likely include Area 5 and MIP. While recording from the contralateral hemisphere might provide more insight into the purely sensory aspects of integrating visual and proprioceptive information, the ipsilateral hemisphere also has purely sensory responses [90, 112, 114, 126, 127] and will provide more insight into how visual and proprioceptive information is integrated for movement.

To the extent that the mechanisms of integration are shared across hemispheres and task conditions, our findings complement previous studies reporting interactions between visual and proprioceptive inputs in several areas of the SPL [35, 90, 126]. These studies reported changes in neurons responses when reaches made with the contralateral hand to visual targets were preformed in the light versus in the dark [90, 126], or when a realistic fake monkey arm was positioned above the contralateral arm in the absence of a task [35]. The second study found that aligning the fake arm with the real arm resulted in a 10% increase in firing rate across the population (at the preferred direction) compared to proprioception alone, which matches the increase observed in our study (Figure 4B&D).

The similarity of these results is striking, both because they were recording from the contralateral hemisphere while we were recording from the ipsilateral hemisphere, and because Graziano et al. found that the visual stimuli had to be an extremely realistic monkey arm to see to this increase, while we saw this increase with a disk at the tip of the fingers. In our study, the monkeys learned that the target disk was coincident with the hand when it was blue and independent of hand position when it was green. This learned visual stimuli modulated neurons responses both on the individual cell and the population level during the task, suggesting that the requirement for realistic vision of the monkey arm may actually be a requirement for relevant sensory stimuli. Further, because we had trials during which the monkey was signaled that proprioceptive information was irrelevant, we were able to compare bimodal responses to responses when vision alone or proprioception alone were the relevant stimuli.

This analysis revealed that changes in bimodal responses relative to cells' best unimodal response follow a much different pattern across the population than changes in response relative to proprioception or vision. We found that the majority of neurons in Area 5 and MIP have bimodal target responses that match their best unimodal responses, showing that much of the difference between proprioceptive or visual responses and bimodal responses is due to heterogeneous task preferences and in these areas. However, while the changes in bimodal versus best unimodal response were quite mixed and only had small effects on the population level, the form of these effects was important in shaping the population responses to bimodal stimuli. For instance, the selective suppression of the minimum bimodal responses across the population during Delay

(Figure 4A&C) contributes to the increased modulation of bimodal responses relative to VIS and PROP. The importance of the selectivity of this suppression is illustrated by the way the general suppression of bimodal relative to the best unimodal responses during Move leads to similar population responses for bimodal and unimodal targets.

The reason for this switch from increased bimodal modulation to similar bimodal and unimodal responses in going from Delay to Move is unclear, but may be related to a shift in the role of the parietal cortex from movement planning to feedback control in these periods [38, 122-125]. If the parietal cortex encodes a forward model of the arm trajectory during movement [122, 124, 125, 137] the strength of the prediction, indicated by the strength of the response, may be less dependent on the sensory input specifying target. However this cannot completely explain our results as responses in MIP are stronger when visual information is available. This suggests that error feedback or feedback control in some areas may split into primarily unisensory streams of information, perhaps to aid in comparing visual and proprioceptive localization [34, 36].

The strong unimodal responses in Area 5 and MIP provide the framework on which the changes in bimodal versus best unimodal responses act to shape population modulation in Delay and Move. The strong responses to both visual and proprioceptive target reaches in Area 5 and MIP are not surprising given the somatosensory and visual inputs to these areas [3, 5, 6, 102, 103] and the visual and proprioceptive responses in these and nearby areas [90, 112, 114, 126, 127]. The importance of strong responses to both sensory modalities for higher bimodal responses across the population is illustrated by the

suppression of bimodal responses in Area 7 (Figure 7), which responds more strongly to visual than proprioceptive inputs [21]. However, it is the fact that that these strong responses to both sensory modalities are occurring in a largely overlapping set of cells that keep this increase from being a trivial summing across different populations. This highlights the importance of the similar tuning properties for visual and proprioceptive responses in individual cells (Chapter 2) in the population encoding of positional information. Representing sensory information in a common form can aid in sensory integration [33, 92, 116, 131], because positional information can be read out by downstream areas independent of it's sensory source. The shared representation of strong visual and proprioceptive responses in Area 5 and MIP provides a population response which can give increased bimodal modulation or similar bimodal and unimodal responses with only small changes in tuning properties across the population (Figure 4 and Figure 8).

In examining multimodal effects in Area 5 and MIP we have focused on how changes individual neuron's responses affect the population response across target types. This is in contrast to previous studies, which used changes in individual neurons bimodal responses relative to their best unimodal responses as the measure of sensory integration [116, 120, 136, 140, 142-145, 150], and leads to very different conclusions about how integration is occurring in Area 5 and MIP. On the single cell level, many cells in Area 5 and MIP show significant, though generally subadditive (Supplemental Figure 1), enhancement or suppression of bimodal relative to the best unimodal responses in both Delay and Move (29% Area 5 Delay, 39% MIP Delay, 41% Area 5 Move, 42% MIP Move). Taken at face

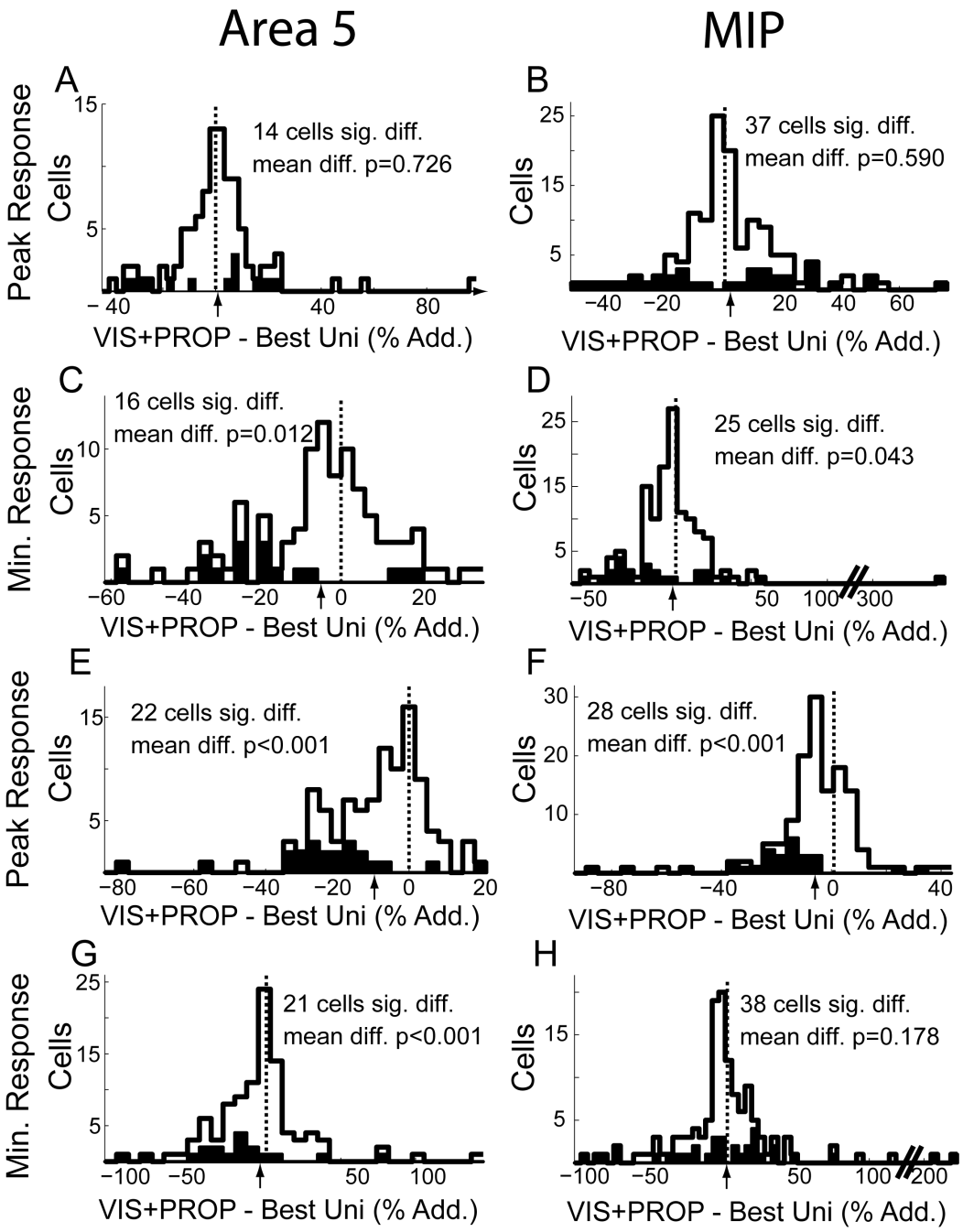
value these changes could indicate that the cells are integrating bimodal sensory information [116, 120, 136, 140, 142-145, 150], though they cannot fit into the framework of additive unimodal responses suggested by network models of sensory integration [82] without some form of response normalization [158]. However, the single cell definition of sensory integration fails to take into account where on the tuning curve these effects are occurring, and how they change the cells modulation, two factors that greatly affect whether the change in a cells response increases or decreases the information encoded by the cell. Several recent studies have looked at how multimodal effects interact with the tuning properties of cells in MST across different sensory inputs [136, 144]. However, when we looked for a similar relationship between the change in bimodal response and correlation in unimodal tuning curves or the predictiveness of a cells firing rate for behavior, we found no relationship between these factors (Figure 6). Thus, to interpret the enhancement and suppression effects of individual cells we needed to look at their affects on tuning properties across the population for a given area. This population analysis revealed increases in the modulation of Delay population responses in Area 5 and MIP that could provide a stronger positional signal about location. These results suggest that integration of visual and proprioceptive information in the SPL may employ different mechanisms of sensory integration than other areas. This difference may be explained by the fact that while other multimodal areas studied likely play a role in representing multimodal information that may or may not share a common source (SC, VIP, MST, [85, 86, 136, 140, 143, 144]), the SPL seems to play a key role in integrating visual and proprioceptive inputs to maintain spatially coherent estimates of arm position for movement [34, 35].

Supplemental Information for Chapter 3: Integration of visual and proprioceptive information in population responses of the posterior parietal cortex.

VIS+PROP versus best unimodal response at peak and minimum

Across both Area 5 and MIP cells showed distinct patterns of enhancement and suppression of bimodal versus unimodal responses at the peak and minimum of the tuning curve. During Delay, cells in both areas had balanced enhancement and suppression of bimodal versus best unimodal responses at the peak of the tuning curve (Supplemental Figure 1A&B). However, at the minimum of the tuning curve, both areas showed slight suppression of bimodal responses, though this suppression only reached significance in Area 5 after correction for multiple comparisons (Supplemental Figure 1C&D, $p=0.012<0.05/2=0.025$ for Area 5, $p=0.043>0.025$ for MIP, paired permutation test, Bonferroni correction). In contrast, during Move, both Area 5 and MIP bimodal responses were suppressed at the peak of the tuning curve, and Area 5 had smaller, though still significant, suppression at the minimum of the tuning curve (Supplemental Figure 1E-H). The majority of the suppression and enhancement effects in both epochs and areas were subadditive, i.e. smaller than the sum of the unimodal responses. The

results suggest that bimodal response modulation is not uniformly expressed across individual cells tuning curves or across behavioral epochs.

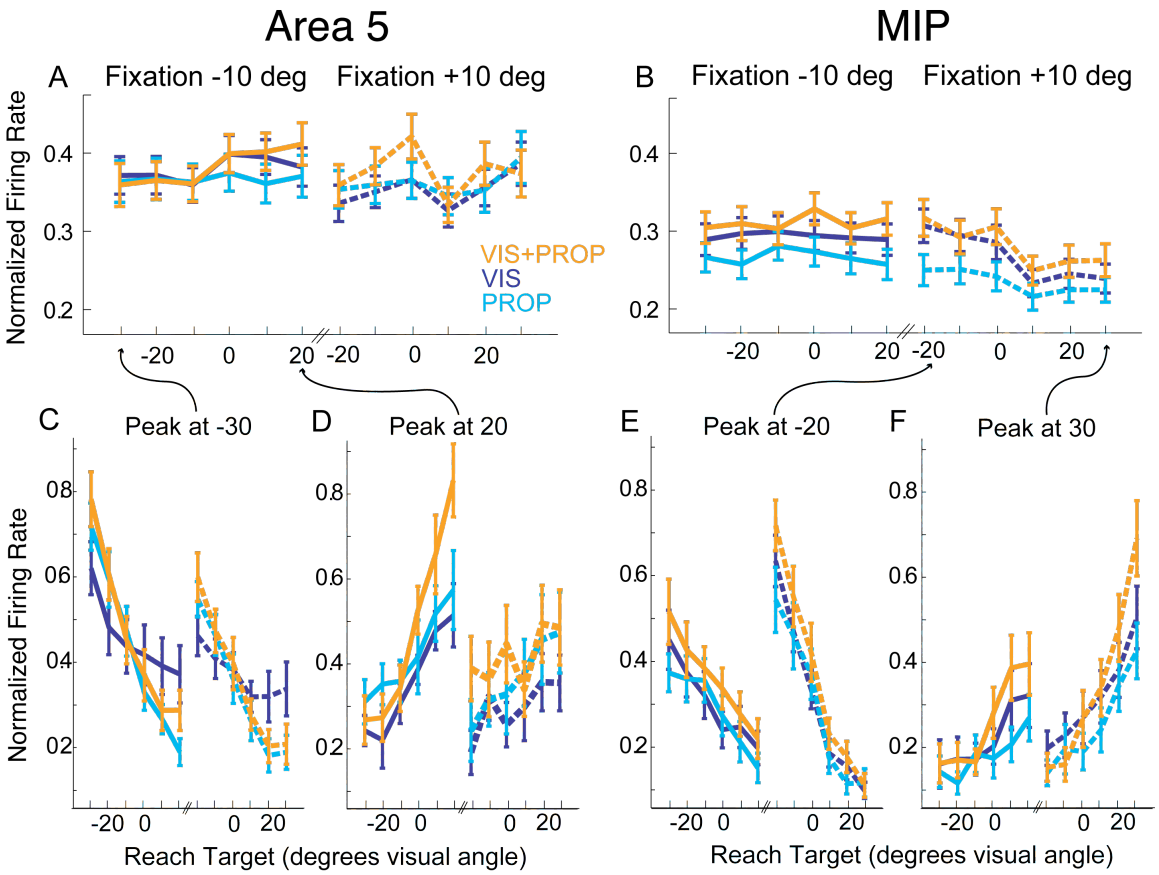


Supplemental Figure 1: Histograms of individual cell differences in VIS+PROP versus best unimodal response at peak and minimum of tuning curve. Differences are divided by the sum of the unimodal responses to give a modulation index (percent of additive response). Filled histograms show cells with significant differences ($p<0.05$, paired permutation test) between

VIS+PROP and best unimodal responses. Arrows indicate mean difference in response. P-values for mean difference come from paired permutation tests, values are significant if $p<0.05/4=0.013$, Bonferroni correction (2 epochs x 2 comparisons).

Delay population responses to VIS+PROP without realignment

The increase in average VIS+PROP responses relative to VIS or PROP responses can be seen even when reach conditions are not reordered to give all cells the same tuning curve. When population responses are averaged across all target and fixation conditions without reordering, VIS+PROP responses are generally greater than or similar to unimodal averages (Supplemental Figure 2A&B). Further, when the responses of small groups of cells with the same preferred direction are averaged, the increase in VIS+PROP responses can still be observed (Supplemental Figure 2C-F). VIS+PROP responses are greater than unimodal responses at the peak in 9/12 Area 5 preferred direction groups and 8/12 MIP preferred direction groups during Delay, with preferred directions with more cells being more likely to show increased responses. The localization of this effect to the peak of the tuning curve is not as clear in these small groups as it is in the reordered population. Nevertheless, the same trends can be seen in small groups of cells selected based on preferred direction as when all cells are reordered. This supports the ideas that: the reordered population averages can be thought of as subsets of cells with the same preferred direction, and changes in specific parts of this reordered population tuning curve are indicative of changes in tuning across the population.

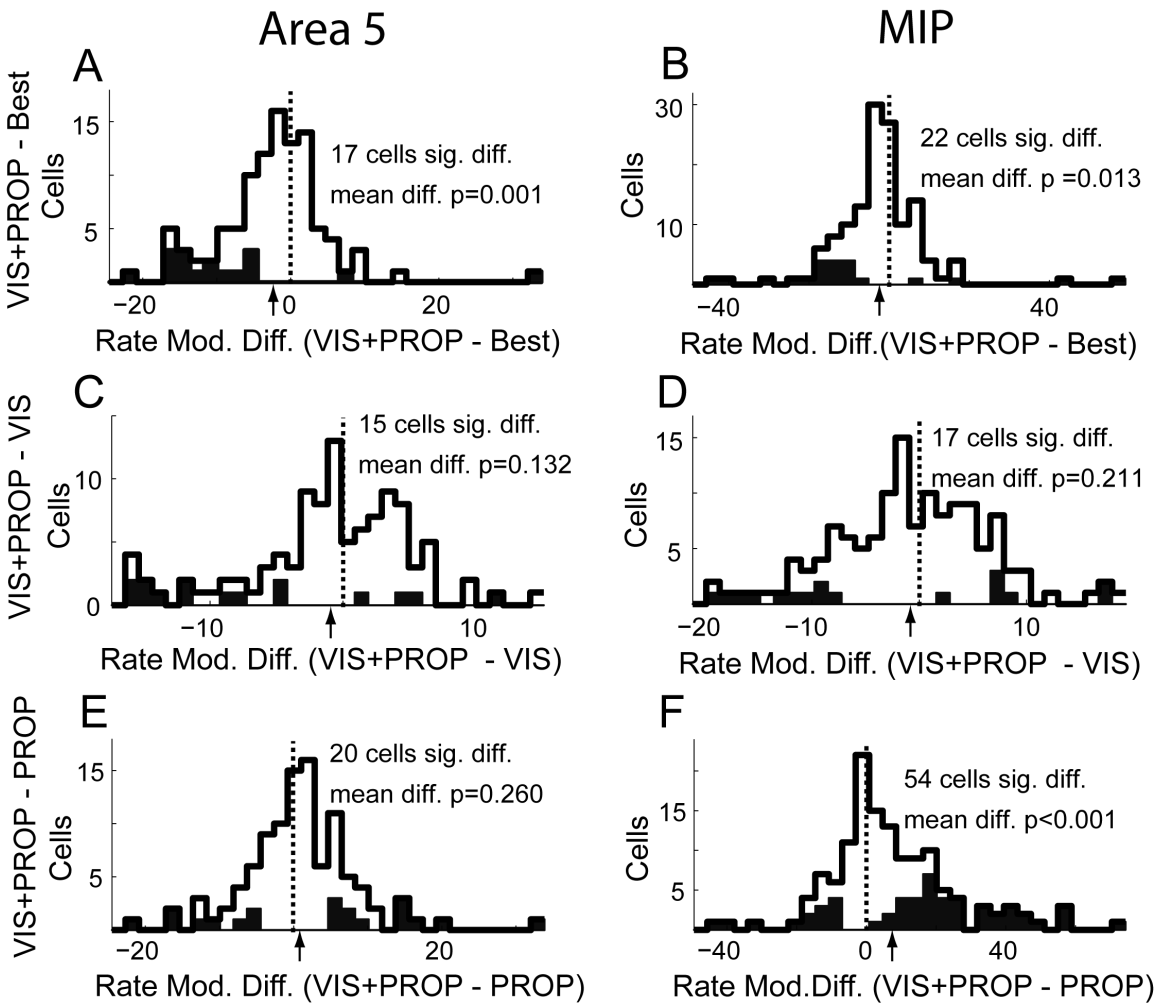


Supplemental Figure 2: Mean normalized population response without reordering. Error bars show standard error. A&B) Show the population average for VIS, VIS+PROP, and PROP for all reach conditions without reordering of conditions to align tuning curves. C-F) Show the average response of all cells that had the highest firing rate at the indicated target. C) Mean across 17 Area 5 cells with peak of tuning at target -30 deg, fixation -10 deg. D) Mean across 8 Area 5 cells with peak of tuning at target +20 deg, fixation -10 deg. E) Mean across 15 MIP cells with peak of tuning at target -20 deg, fixation +10 deg. F) Mean across 14 MIP cells with peak of tuning at target +30 deg, fixation +10 deg.

Move changes in individual cells VIS+PROP modulation

Individual cells in Area 5 and MIP show decreased modulation of VIS+PROP versus best unimodal responses (Supplemental Figure 3A&B), but similar modulation across the population for VIS+PROP and unimodal responses, during Move (Supplemental Figure 3C-F, similar to VIS and PROP in Area 5, similar to VIS in MIP).

This suggests that the suppression of individual neurons bimodal versus best unimodal responses during Move serves to equalize the overall bimodal and unimodal responses in these areas.



Supplemental Figure 3: Histograms of individual cell differences in Move modulation for Area 5 and MIP. A&B) Differences between VIS+PROP and each cells best unimodal response modulation. C&D) Differences between cells VIS+PROP and VIS modulation. E&F) Differences between cells VIS+PROP and PROP modulation. Filled histograms show individual cells with significant differences. Arrows indicate mean difference in response. P-values for mean difference come from paired permutation tests.

Literature Cited

1. Kandel, E.R., Schwartz, J.H. & Jessell, T.M. Principles of Neural Science. (2000).
2. Andersen, R.A. & Buneo, C.A. Sensorimotor integration in posterior parietal cortex. *Adv Neurol* **93**, 159-177 (2003).
3. Kalaska, J.F. Parietal cortex area 5 and visuomotor behavior. *Can J Physiol Pharmacol* **74**, 483-498 (1996).
4. Kalaska, J.F., Scott, S.H., Cisek, P. & Sergio, L.E. Cortical control of reaching movements. *Curr Opin Neurobiol* **7**, 849-859 (1997).
5. Rizzolatti, G., Fogassi, L. & Gallese, V. Parietal cortex: from sight to action. *Curr Opin Neurobiol* **7**, 562-567 (1997).
6. Wise, S.P., Boussaoud, D., Johnson, P.B. & Caminiti, R. Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annu Rev Neurosci* **20**, 25-42 (1997).
7. Caminiti, R., Ferraina, S. & Johnson, P.B. The sources of visual information to the primate frontal lobe: a novel role for the superior parietal lobule. *Cereb Cortex* **6**, 319-328 (1996).
8. Batista, A.P., Buneo, C.A., Snyder, L.H. & Andersen, R.A. Reach plans in eye-centered coordinates. *Science* **285**, 257-260 (1999).
9. Buneo, C.A., Batista, A.P., Jarvis, M.R. & Andersen, R.A. Time-invariant reference frames for parietal reach activity. *Exp Brain Res* **188**, 77-89 (2008).
10. Buneo, C.A., Jarvis, M.R., Batista, A.P. & Andersen, R.A. Direct visuomotor transformations for reaching. *Nature* **416**, 632-636 (2002).
11. Chang, S.W. & Snyder, L.H. Idiosyncratic and systematic aspects of spatial representations in the macaque parietal cortex. *Proc Natl Acad Sci U S A* (2010).
12. Ferraina, S.*, et al.* Reaching in depth: hand position dominates over binocular eye position in the rostral superior parietal lobule. *J Neurosci* **29**, 11461-11470 (2009).
13. Lacquaniti, F., Guigon, E., Bianchi, L., Ferraina, S. & Caminiti, R. Representing spatial information for limb movement: role of area 5 in the monkey. *Cereb Cortex* **5**, 391-409 (1995).
14. Fattori, P., Kutz, D.F., Breveglieri, R., Marzocchi, N. & Galletti, C. Spatial tuning of reaching activity in the medial parieto-occipital cortex (area V6A) of macaque monkey. *Eur J Neurosci* **22**, 956-972 (2005).
15. Cohen, Y.E. & Andersen, R.A. Reaches to sounds encoded in an eye-centered reference frame. *Neuron* **27**, 647-652 (2000).
16. Mullette-Gillman, O.A., Cohen, Y.E. & Groh, J.M. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J Neurophysiol* **94**, 2331-2352 (2005).

17. Stricanne, B., Andersen, R.A. & Mazzoni, P. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J Neurophysiol* **76**, 2071-2076 (1996).
18. Batista, A.P.*, et al.* Reference frames for reach planning in macaque dorsal premotor cortex. *J Neurophysiol* **98**, 966-983 (2007).
19. Burnod, Y.*, et al.* Parieto-frontal coding of reaching: an integrated framework. *Exp Brain Res* **129**, 325-346 (1999).
20. Caminiti, R., Ferraina, S. & Mayer, A.B. Visuomotor transformations: early cortical mechanisms of reaching. *Curr Opin Neurobiol* **8**, 753-761 (1998).
21. Colby, C.L. & Duhamel, J.R. Spatial representations for action in parietal cortex. *Brain Res Cogn Brain Res* **5**, 105-115 (1996).
22. Graziano, M.S. Is reaching eye-centered, body-centered, hand-centered, or a combination? *Rev Neurosci* **12**, 175-185 (2001).
23. Mullette-Gillman, O.A., Cohen, Y.E. & Groh, J.M. Motor-Related Signals in the Intraparietal Cortex Encode Locations in a Hybrid, rather than Eye-Centered Reference Frame. *Cereb Cortex* (2008).
24. Pesaran, B., Nelson, M.J. & Andersen, R.A. Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning. *Neuron* **51**, 125-134 (2006).
25. Wu, W. & Hatsopoulos, N. Evidence against a single coordinate system representation in the motor cortex. *Exp Brain Res* **175**, 197-210 (2006).
26. Wu, W. & Hatsopoulos, N.G. Coordinate system representations of movement direction in the premotor cortex. *Exp Brain Res* **176**, 652-657 (2007).
27. Avillac, M., Deneve, S., Olivier, E., Pouget, A. & Duhamel, J.R. Reference frames for representing visual and tactile locations in parietal cortex. *Nat Neurosci* **8**, 941-949 (2005).
28. Fetsch, C.R., Wang, S., Gu, Y., Deangelis, G.C. & Angelaki, D.E. Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area. *J Neurosci* **27**, 700-712 (2007).
29. Jay, M.F. & Sparks, D.L. Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J Neurophysiol* **57**, 35-55 (1987).
30. Sober, S.J. & Sabes, P.N. Multisensory integration during motor planning. *J Neurosci* **23**, 6982-6992 (2003).
31. Sober, S.J. & Sabes, P.N. Flexible strategies for sensory integration during motor planning. *Nat Neurosci* **8**, 490-497 (2005).
32. Medendorp, W.P., Goltz, H.C., Vilis, T. & Crawford, J.D. Gaze-centered updating of visual space in human parietal cortex. *J Neurosci* **23**, 6209-6214 (2003).
33. Cohen, Y.E. & Andersen, R.A. A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci* **3**, 553-562 (2002).
34. Clower, D.M.*, et al.* Role of posterior parietal cortex in the recalibration of visually guided reaching. *Nature* **383**, 618-621 (1996).
35. Graziano, M.S., Cooke, D.F. & Taylor, C.S. Coding the location of the arm by sight. *Science* **290**, 1782-1786 (2000).
36. Simani, M.C., McGuire, L.M. & Sabes, P.N. Visual-shift adaptation is composed of separable sensory and task-dependent effects. *J Neurophysiol* **98**, 2827-2841 (2007).

37. Schlicht, E.J. & Schrater, P.R. Impact of coordinate transformation uncertainty on human sensorimotor control. *J Neurophysiol* **97**, 4203-4214 (2007).
38. Buneo, C.A. & Andersen, R.A. The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia* **44**, 2594-2606 (2006).
39. Flanders, M., Helms Tillery, S.I. & Soechting, J.F. Early stages in a sensorimotor transformation. *Behav Brain Sci* **15**, 309-362 (1992).
40. Lacquaniti, F. & Caminiti, R. Visuo-motor transformations for arm reaching. *Eur J Neurosci* **10**, 195-203 (1998).
41. Soechting, J.F. & Flanders, M. Moving in three-dimensional space: frames of reference, vectors, and coordinate systems. *Annu Rev Neurosci* **15**, 167-191 (1992).
42. Beurze, S.M., Van Pelt, S. & Medendorp, W.P. Behavioral reference frames for planning human reaching movements. *J Neurophysiol* **96**, 352-362 (2006).
43. Engel, K.C., Flanders, M. & Soechting, J.F. Oculocentric frames of reference for limb movement. *Archives italiennes de biologie* **140**, 211-219 (2002).
44. Henriques, D.Y., Klier, E.M., Smith, M.A., Lowy, D. & Crawford, J.D. Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci* **18**, 1583-1594 (1998).
45. Pouget, A., Ducom, J.C., Torri, J. & Bavelier, D. Multisensory spatial representations in eye-centered coordinates for reaching. *Cognition* **83**, B1-11 (2002).
46. McIntyre, J., Stratta, F. & Lacquaniti, F. Viewer-centered frame of reference for pointing to memorized targets in three-dimensional space. *J Neurophysiol* **78**, 1601-1618 (1997).
47. Carrozzo, M., McIntyre, J., Zago, M. & Lacquaniti, F. Viewer-centered and body-centered frames of reference in direct visuomotor transformations. *Exp Brain Res* **129**, 201-210 (1999).
48. McIntyre, J., Stratta, F. & Lacquaniti, F. Short-term memory for reaching to visual targets: psychophysical evidence for body-centered reference frames. *J Neurosci* **18**, 8423-8435 (1998).
49. Carrozzo, M. & Lacquaniti, F. A hybrid frame of reference for visuo-manual coordination. *Neuroreport* **5**, 453-456 (1994).
50. Bock, O. Contribution of retinal versus extraretinal signals towards visual localization in goal-directed movements. *Exp Brain Res* **64**, 476-482 (1986).
51. Bock, O. Localization of objects in the peripheral visual field. *Behav Brain Res* **56**, 77-84 (1993).
52. Sorrento, G.U. & Henriques, D.Y. Reference frame conversions for repeated arm movements. *J Neurophysiol* (2008).
53. Soechting, J.F. & Flanders, M. Sensorimotor representations for pointing to targets in three-dimensional space. *J Neurophysiol* **62**, 582-594 (1989).
54. Van Pelt, S. & Medendorp, W.P. Updating target distance across eye movements in depth. *J Neurophysiol* **99**, 2281-2290 (2008).
55. Vetter, P., Goodbody, S.J. & Wolpert, D.M. Evidence for an eye-centered spherical representation of the visuomotor map. *J Neurophysiol* **81**, 935-939 (1999).
56. Vindras, P. & Viviani, P. Frames of reference and control parameters in visuomanual pointing. *J Exp Psychol Hum Percept Perform* **24**, 569-591 (1998).

57. Knill, D.C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci* **27**, 712-719 (2004).
58. Blohm, G. & Crawford, J.D. Computations for geometrically accurate visually guided reaching in 3-D space. *J Vis* **7**, 4 1-22 (2007).
59. Land, M.F. & Hayhoe, M. In what ways do eye movements contribute to everyday activities? *Vision Res* **41**, 3559-3565 (2001).
60. van Beers, R.J., Sittig, A.C. & Denier van der Gon, J.J. The precision of proprioceptive position sense. *Exp Brain Res* **122**, 367-377 (1998).
61. Neggers, S.F. & Bekkering, H. Gaze anchoring to a pointing target is present during the entire pointing movement and is driven by a non-visual signal. *J Neurophysiol* **86**, 961-970 (2001).
62. Neggers, S.F. & Bekkering, H. Coordinated control of eye and hand movements in dynamic reaching. *Hum Mov Sci* **21**, 349-376 (2002).
63. Bayes, T. An essay towards solving a problem in the doctrine of chances. 1763. *MD Comput* **8**, 157-171 (1991).
64. Balslev, D. & Miall, R.C. Eye position representation in human anterior parietal cortex. *J Neurosci* **28**, 8968-8972 (2008).
65. Ernst, M.O. & Banks, M.S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429-433 (2002).
66. Ernst, M.O. & Bulthoff, H.H. Merging the senses into a robust percept. *Trends Cogn Sci* **8**, 162-169 (2004).
67. Helbig, H.B. & Ernst, M.O. Optimal integration of shape information from vision and touch. *Exp Brain Res* **179**, 595-606 (2007).
68. Knill, D.C. Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J Vis* **7**, 5 1-24 (2007).
69. Kording, K.P. & Wolpert, D.M. Probabilistic mechanisms in sensorimotor control. *Novartis Found Symp* **270**, 191-198; discussion 198-202, 232-197 (2006).
70. van Beers, R.J., Sittig, A.C. & Gon, J.J. Integration of proprioceptive and visual position-information: An experimentally supported model. *J Neurophysiol* **81**, 1355-1364 (1999).
71. Yuille, A.L. & Bulthoff, H.H. Bayesian decision theory and psychophysics. in *Perception as Bayesian Inference* 123-136 (Published by Cambridge University Press, 1996).
72. Enright, J.T. The non-visual impact of eye orientation on eye-hand coordination. *Vision Res* **35**, 1611-1618 (1995).
73. Colby, C.L., Duhamel, J.R. & Goldberg, M.E. Oculocentric spatial representation in parietal cortex. *Cereb Cortex* **5**, 470-481 (1995).
74. Duhamel, J.R., Colby, C.L. & Goldberg, M.E. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* **255**, 90-92 (1992).
75. Merriam, E.P., Genovese, C.R. & Colby, C.L. Spatial updating in human parietal cortex. *Neuron* **39**, 361-373 (2003).
76. Henriques, D.Y. & Crawford, J.D. Direction-dependent distortions of retinocentric space in the visuomotor transformation for pointing. *Exp Brain Res* **132**, 179-194 (2000).

77. Henriques, D.Y. & Crawford, J.D. Role of eye, head, and shoulder geometry in the planning of accurate arm movements. *J Neurophysiol* **87**, 1677-1685 (2002).
78. Henriques, D.Y., Medendorp, W.P., Gielen, C.C. & Crawford, J.D. Geometric computations underlying eye-hand coordination: orientations of the two eyes and the head. *Exp Brain Res* **152**, 70-78 (2003).
79. Lewald, J. & Ehrenstein, W.H. Visual and proprioceptive shifts in perceived egocentric direction induced by eye-position. *Vision Res* **40**, 539-547 (2000).
80. Vindras, P., Desmurget, M., Prablanc, C. & Viviani, P. Pointing errors reflect biases in the perception of the initial hand position. *J Neurophysiol* **79**, 3290-3294 (1998).
81. Soechting, J.F. & Flanders, M. Errors in pointing are due to approximations in sensorimotor transformations. *J Neurophysiol* **62**, 595-608 (1989).
82. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat Neurosci* **9**, 1432-1438 (2006).
83. Shadlen, M.N. & Newsome, W.T. Noise, neural codes and cortical organization. *Curr Opin Neurobiol* **4**, 569-579 (1994).
84. Helbig, H.B. & Ernst, M.O. Knowledge about a common source can promote visual- haptic integration. *Perception* **36**, 1523-1533 (2007).
85. Kording, K.P.*, et al.* Causal inference in multisensory perception. *PLoS ONE* **2**, e943 (2007).
86. Kording, K.P. & Tenenbaum, J.B. Causal inference in multisensory integration. in *NIPS* 737-744 (MIT Press, 2006).
87. Ahmed, A.A., Wolpert, D.M. & Flanagan, J.R. Flexible representations of dynamics are used in object manipulation. *Curr Biol* **18**, 763-768 (2008).
88. Kluzik, J., Diedrichsen, J., Shadmehr, R. & Bastian, A.J. Reach adaptation: what determines whether we learn an internal model of the tool or adapt the model of our arm? *J Neurophysiol* **100**, 1455-1464 (2008).
89. Crawford, J.D., Medendorp, W.P. & Marotta, J.J. Spatial transformations for eye-hand coordination. *J Neurophysiol* **92**, 10-19 (2004).
90. Battaglia-Mayer, A.*, et al.* Eye-hand coordination during reaching. II. An analysis of the relationships between visuomanual signals in parietal cortex and parieto-frontal association projections. *Cereb Cortex* **11**, 528-544 (2001).
91. Caminiti, R.*, et al.* Early coding of reaching: frontal and parietal association connections of parieto-occipital cortex. *Eur J Neurosci* **11**, 3339-3345 (1999).
92. Battaglia-Mayer, A., Caminiti, R., Lacquaniti, F. & Zago, M. Multiple levels of representation of reaching in the parieto-frontal network. *Cereb Cortex* **13**, 1009-1022 (2003).
93. Good, P.I. Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, 2nd ed. (2000).
94. Gordon, J., Ghilardi, M.F., Cooper, S.E. & Ghez, C. Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. *Exp Brain Res* **99**, 112-130 (1994).
95. Scott, S.H. & Loeb, G.E. The computation of position sense from spindles in mono- and multiarticular muscles. *J Neurosci* **14**, 7529-7540 (1994).
96. van Opstal, A.J. & van Gisbergen, J.A. Scatter in the metrics of saccades and properties of the collicular motor map. *Vision Res* **29**, 1183-1196 (1989).

97. Deneve, S., Latham, P.E. & Pouget, A. Efficient computation and cue integration with noisy population codes. *Nat Neurosci* **4**, 826-831 (2001).
98. Pratt, J. & Turk-Browne, N.B. The attentional repulsion effect in perception and action. *Exp Brain Res* **152**, 376-382 (2003).
99. Suzuki, S. & Cavanagh, P. Focused attention distorts visual space: an attentional repulsion effect. *J Exp Psychol Hum Percept Perform* **23**, 443-463 (1997).
100. Neggers, S.F. & Bekkering, H. Ocular gaze is anchored to the target of an ongoing pointing movement. *J Neurophysiol* **83**, 639-651 (2000).
101. Musseler, J., van der Heijden, A.H., Mahmud, S.H., Deubel, H. & Ertsey, S. Relative mislocalization of briefly presented stimuli in the retinal periphery. *Percept Psychophys* **61**, 1646-1661 (1999).
102. Johnson, P.B., Ferraina, S., Bianchi, L. & Caminiti, R. Cortical networks for visual reaching: physiological and anatomical organization of frontal and parietal lobe arm regions. *Cereb Cortex* **6**, 102-119 (1996).
103. Marconi, B*., et al.* Eye-hand coordination during reaching. I. Anatomical relationships between parietal and frontal cortex. *Cereb Cortex* **11**, 513-527 (2001).
104. Pandya, D.N. & Seltzer, B. Intrinsic connections and architectonics of posterior parietal cortex in the rhesus monkey. *J Comp Neurol* **204**, 196-210 (1982).
105. Bakola, S., Gamberini, M., Passarelli, L., Fattori, P. & Galletti, C. Cortical Connections of Parietal Field PEc in the Macaque: Linking Vision and Somatic Sensation for the Control of Limb Action. *Cereb Cortex* (2010).
106. Galletti, C., Kutz, D.F., Gamberini, M., Breveglieri, R. & Fattori, P. Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Exp Brain Res* **153**, 158-170 (2003).
107. McGuire, L.M. & Sabes, P.N. Sensory transformations and the use of multiple reference frames for reach planning. *Nat Neurosci* **12**, 1056-1061 (2009).
108. Soechting, J.F. & Flanders, M. Arm movements in three-dimensional space: computation, theory, and observation. *Exerc Sport Sci Rev* **19**, 389-418 (1991).
109. Deneve, S. & Pouget, A. Bayesian multisensory integration and cross-modal spatial links. *J Physiol Paris* **98**, 249-258 (2004).
110. Pouget, A., Deneve, S. & Duhamel, J.R. A computational perspective on the neural basis of multisensory spatial representations. *Nat Rev Neurosci* **3**, 741-747 (2002).
111. Mountcastle, V.B., Lynch, J.C., Georgopoulos, A., Sakata, H. & Acuna, C. Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *J Neurophysiol* **38**, 871-908 (1975).
112. Breveglieri, R., Galletti, C., Monaco, S. & Fattori, P. Visual, somatosensory, and bimodal activities in the macaque parietal area PEc. *Cereb Cortex* **18**, 806-816 (2008).
113. Sakata, H., Takaoka, Y., Kawarasaki, A. & Shibutani, H. Somatosensory properties of neurons in the superior parietal cortex (area 5) of the rhesus monkey. *Brain research* **64**, 85-102 (1973).
114. Breveglieri, R., Kutz, D.F., Fattori, P., Gamberini, M. & Galletti, C. Somatosensory cells in the parieto-occipital area V6A of the macaque. *Neuroreport* **13**, 2113-2116 (2002).

115. Galletti, C.*, et al.* The cortical connections of area V6: an occipito-parietal network processing visual information. *Eur J Neurosci* **13**, 1572-1588 (2001).
116. Stein, B.E. & Stanford, T.R. Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* **9**, 255-266 (2008).
117. Snyder, L.H., Batista, A.P. & Andersen, R.A. Coding of intention in the posterior parietal cortex. *Nature* **386**, 167-170 (1997).
118. Efron, B. & Tibshirani, R.J. An introduction to the bootstrap. (1993).
119. Andersen, R.A., Snyder, L.H., Batista, A.P., Buneo, C.A. & Cohen, Y.E. Posterior parietal areas specialized for eye movements (LIP) and reach (PRR) using a common coordinate frame. *Novartis Found Symp* **218**, 109-122; discussion 122-108, 171-105 (1998).
120. Angelaki, D.E., Gu, Y. & DeAngelis, G.C. Multisensory integration: psychophysics, neurophysiology, and computation. *Curr Opin Neurobiol* **19**, 452-458 (2009).
121. Cohen, Y.E. Multimodal activity in the parietal cortex. *Hear Res* **258**, 100-105 (2009).
122. Archambault, P.S., Caminiti, R. & Battaglia-Mayer, A. Cortical mechanisms for online control of hand movement trajectory: the role of the posterior parietal cortex. *Cereb Cortex* **19**, 2848-2864 (2009).
123. Desmurget, M.*, et al.* Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nat Neurosci* **2**, 563-567 (1999).
124. Desmurget, M. & Grafton, S. Forward modeling allows feedback control for fast reaching movements. *Trends Cogn Sci* **4**, 423-431 (2000).
125. Mulliken, G.H., Musallam, S. & Andersen, R.A. Forward estimation of movement state in posterior parietal cortex. *Proc Natl Acad Sci U S A* **105**, 8170-8177 (2008).
126. Battaglia-Mayer, A.*, et al.* Early coding of reaching in the parietooccipital cortex. *J Neurophysiol* **83**, 2374-2391 (2000).
127. Breveglieri, R., Galletti, C., Gamberini, M., Passarelli, L. & Fattori, P. Somatosensory cells in area PEc of macaque posterior parietal cortex. *J Neurosci* **26**, 3679-3684 (2006).
128. Chang, S.W., Dickinson, A.R. & Snyder, L.H. Limb-specific representation for reaching in the posterior parietal cortex. *J Neurosci* **28**, 6128-6140 (2008).
129. Schiller, P.H. & Tehovnik, E.J. Neural mechanisms underlying target selection with saccadic eye movements. *Prog Brain Res* **149**, 157-171 (2005).
130. Groh, J.M. & Sparks, D.L. Saccades to somatosensory targets. III. eye-position-dependent somatosensory activity in primate superior colliculus. *J Neurophysiol* **75**, 439-453 (1996).
131. Battaglia-Mayer, A. & Caminiti, R. Optic ataxia as a result of the breakdown of the global tuning fields of parietal neurones. *Brain* **125**, 225-237 (2002).
132. Beurze, S.M., de Lange, F.P., Toni, I. & Medendorp, W.P. Integration of target and effector information in the human brain during reach planning. *J Neurophysiol* **97**, 188-199 (2007).
133. Grefkes, C. & Fink, G.R. The functional organization of the intraparietal sulcus in humans and monkeys. *J Anat* **207**, 3-17 (2005).

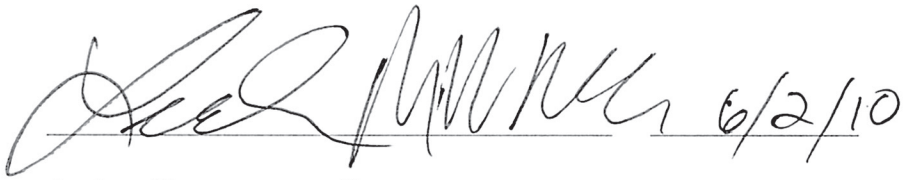
134. Battaglia-Mayer, A., Archambault, P.S. & Caminiti, R. The cortical network for eye-hand coordination and its relevance to understanding motor disorders of parietal patients. *Neuropsychologia* **44**, 2607-2620 (2006).
135. Mascaro, M., Battaglia-Mayer, A., Nasi, L., Amit, D.J. & Caminiti, R. The eye and the hand: neural mechanisms and network models for oculomanual coordination in parietal cortex. *Cereb Cortex* **13**, 1276-1286 (2003).
136. Gu, Y., Angelaki, D.E. & Deangelis, G.C. Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci* **11**, 1201-1210 (2008).
137. Andersen, R.A. & Cui, H. Intention, action planning, and decision making in parietal-frontal circuits. *Neuron* **63**, 568-583 (2009).
138. Cohen, Y.E., Batista, A.P. & Andersen, R.A. Comparison of neural activity preceding reaches to auditory and visual stimuli in the parietal reach region. *Neuroreport* **13**, 891-894 (2002).
139. Burr, D. & Alais, D. Combining visual and auditory information. *Prog Brain Res* **155**, 243-258 (2006).
140. Meredith, M.A. & Stein, B.E. Interactions among converging sensory inputs in the superior colliculus. *Science* **221**, 389-391 (1983).
141. Rowland, B.A. & Stein, B.E. Temporal profiles of response enhancement in multisensory integration. *Front Neurosci* **2**, 218-224 (2008).
142. Stanford, T.R. & Stein, B.E. Superadditivity in multisensory integration: putting the computation in context. *Neuroreport* **18**, 787-792 (2007).
143. Avillac, M., Ben Hamed, S. & Duhamel, J.R. Multisensory integration in the ventral intraparietal area of the macaque monkey. *J Neurosci* **27**, 1922-1932 (2007).
144. Morgan, M.L., Deangelis, G.C. & Angelaki, D.E. Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* **59**, 662-673 (2008).
145. Sugihara, T., Diltz, M.D., Averbeck, B.B. & Romanski, L.M. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *J Neurosci* **26**, 11138-11147 (2006).
146. Andersen, R.A. Multimodal integration for the representation of space in the posterior parietal cortex. *Philos Trans R Soc Lond B Biol Sci* **352**, 1421-1428 (1997).
147. Fattori, P., Gamberini, M., Kutz, D.F. & Galletti, C. 'Arm-reaching' neurons in the parietal area V6A of the macaque monkey. *Eur J Neurosci* **13**, 2309-2313 (2001).
148. Taira, M., Mine, S., Georgopoulos, A.P., Murata, A. & Sakata, H. Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp Brain Res* **83**, 29-36 (1990).
149. Johnson, P.B., Ferraina, S. & Caminiti, R. Cortical networks for visual reaching. *Exp Brain Res* **97**, 361-365 (1993).
150. Stein, B.E., Stanford, T.R., Ramachandran, R., Perrault, T.J., Jr. & Rowland, B.A. Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Exp Brain Res* **198**, 113-126 (2009).
151. Iwamura, Y., Iriki, A. & Tanaka, M. Bilateral hand representation in the postcentral somatosensory cortex. *Nature* **369**, 554-556 (1994).
152. Iwamura, Y., Tanaka, M., Iriki, A., Taoka, M. & Toda, T. Processing of tactile and kinesthetic signals from bilateral sides of the body in the postcentral gyrus of awake monkeys. *Behav Brain Res* **135**, 185-190 (2002).

153. Cisek, P., Crammond, D.J. & Kalaska, J.F. Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm. *J Neurophysiol* **89**, 922-942 (2003).
154. Donchin, O., Gribova, A., Steinberg, O., Bergman, H. & Vaadia, E. Primary motor cortex is involved in bimanual coordination. *Nature* **395**, 274-278 (1998).
155. Donchin, O.*, et al.* Single-unit activity related to bimanual arm movements in the primary and supplementary motor cortices. *J Neurophysiol* **88**, 3498-3517 (2002).
156. Tanji, J., Okano, K. & Sato, K.C. Relation of neurons in the nonprimary motor cortex to bilateral hand movement. *Nature* **327**, 618-620 (1987).
157. Tanji, J., Okano, K. & Sato, K.C. Neuronal activity in cortical motor areas related to ipsilateral, contralateral, and bilateral digit movements of the monkey. *J Neurophysiol* **60**, 325-343 (1988).
158. Ma, W.J. & Pouget, A. Linking neurons to behavior in multisensory perception: a computational review. *Brain research* **1242**, 4-12 (2008).

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature Date