

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

Leah M M McGuire & Philip N Sabes

The sensory signals that drive movement planning arrive in a variety of 'reference frames', and integrating or comparing them requires sensory transformations. We propose a model in which the statistical properties of sensory signals and their transformations determine how these signals are used. This model incorporates the patterns of gaze-dependent errors that we found in our human psychophysics experiment when the sensory signals available for reach planning were varied. These results challenge the widely held ideas that error patterns directly reflect the reference frame of the underlying neural representation and that it is preferable to use a single common reference frame for movement planning. We found that gaze-dependent error patterns, often cited as evidence for retinotopic reach planning, can be explained by a transformation bias and are not exclusively linked to retinotopic representations. Furthermore, the presence of multiple reference frames allows for optimal use of available sensory information and explains task-dependent reweighting of sensory signals.

Humans use various sensory signals when interacting with the environment. We can reach to pick up a coin that we see in front of us or transfer the coin from one hand to another without looking. Using multiple sensory modalities for planning similar movements is potentially problematic, as different sensory signals arrive in different reference frames. Specifically, early visual pathways represent stimulus location relative to current gaze location (a retinotopic representation), whereas proprioceptive signals represent hand location relative to the shoulder or trunk (a body-centered representation). To utilize these signals, some of them must be transformed between reference frames. Although sensory transformations may appear mathematically simple, we found that transformations can incur a cost by adding bias and variability^{1,2} into the transformed signal. Thus, sensory transformations likely influence the flow of information in motor planning circuits.

It has been argued that transforming sensory signals into a common representation would simplify reach planning³⁻⁷, and many researchers have attempted to characterize this representation. Many psychophysical studies have focused on the patterns of reach error using the assumption that the reference frame for movement planning directly determines the spatial pattern of errors. Such studies argue for both retinotopic^{5,8-11} and hand- or body-centered^{12,13} planning. Studies of primate physiology and human functional magnetic resonance imaging have also found evidence for a range of neural representations for movement planning¹⁴⁻²³. These disparate results suggest that a single common representation for reach planning is unlikely²²⁻²⁶. We argue that the presence of noisy sensory transformations makes it advantageous to represent movement plans simultaneously in multiple reference frames.

We examined how the representations of movement plan depend on the available sensory inputs, focusing on the effects of gaze location.

A well-studied gaze-dependent error is the retinal eccentricity effect, where subjects overestimate the distance between the gaze location and a visually peripheral target when pointing to the target²⁷. Because these errors are most parsimoniously described as an overshoot in a retinotopic reference frame, they are cited as evidence for retinotopic reach planning⁸⁻¹⁰. We found that the magnitude and direction of gaze-dependent errors depended heavily on the available sensory information. These results were interpreted using a model of movement planning in which sensory signals are combined in a statistically principled manner in two separate reference frames. Our model provides an explanation for these gaze-dependent reach errors: they arise when sensory information about target location is transformed between representations using an internal estimate of gaze direction that is biased toward the target. We thus demonstrate that spatial patterns of reach errors do not necessarily directly 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 that are needed to compute a movement vector. Target information was varied by having subjects reach to either visual targets (VIS), proprioceptive targets (the left index finger, PROP) or targets consisting of simultaneous visual and proprioceptive signals (left index finger with visual feedback, VIS + PROP). Information about initial hand position was varied by having subjects reach either with (feedback) or without

W. M. Keck Center for Integrative Neuroscience, Department of Physiology, and the Neuroscience Graduate Program, University of California, San Francisco, California, USA. Correspondence should be addressed to P.N.S. (sabes@phy.ucsf.edu).

Received 23 March; accepted 1 June; published online 13 July 2009; doi:10.1038/nn.2357

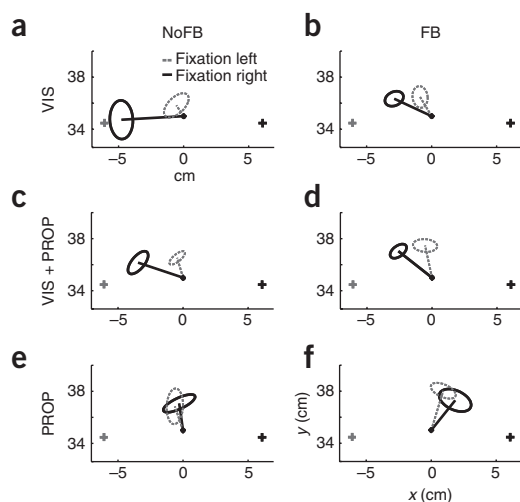


Figure 1 Reach errors at the center target for an example subject for all trial types. (a–f) Lines indicate mean reach error for each gaze position, ellipses represent s.d., + indicates fixation points and • indicates reach targets. The origin (not shown) is located directly below the midpoint of the eyes. FB, feedback trial.

(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, we measured movement errors as subjects reached to nine different target locations with gaze held on one of two fixation points (**Supplementary Fig. 1**).

A comparison of reach endpoints for an example subject at the midline target shows how reach errors depended on the available sensory signals (**Fig. 1**). The mean error changed markedly as a function of gaze location and these gaze-dependent effects differed across trial types. During VIS and VIS + PROP trials (**Fig. 1a–d**), reach endpoints were biased away from the gaze location (the retinal eccentricity effect) and the magnitude of the effect decreased with increasing sensory information (compare **Figs. 1a** and **1b–d**). In PROP trials, a small bias toward gaze location was observed instead (**Fig. 1e,f**). These patterns were consistent across targets (**Supplementary Section 1** and **Supplementary Fig. 2**). In addition to these gaze-dependent effects, there was a gaze-independent bias in reaching that could differ across targets and trial types. Although there was a trend toward overshooting the target, the pattern of this bias, measured across targets and trial types in a separate gaze-free trial condition, was idiosyncratic from subject to subject (**Supplementary Section 1** and **Supplementary Fig. 3**), making these patterns difficult to interpret. We therefore focused on the consistent gaze-dependent effects.

To isolate the gaze-dependent effects, we analyzed reach errors in polar coordinates about the midpoint of the eyes and subtracted the gaze-free errors (Online Methods, **Supplementary Section 1** and **Supplementary Fig. 4**). When we examined radial (depth) reach errors, we found a tendency to overshoot the target; however, these errors did not differ between the two gaze locations (**Supplementary Section 1** and **Supplementary Fig. 5**). When we examined angular reach errors (**Fig. 2**), we observed both a general leftward bias (see **Supplementary Section 1**) and significant differences as a function of gaze location ($P < 0.001$ and $P = 0.003$). Gaze-dependent effects were thus confined to the angular reach errors.

We next examined how these gaze-dependent effects varied with trial type (**Fig. 2a–f**). The retinal eccentricity effect was observed in VIS/NoFB trials (**Fig. 2a**); subjects made rightward (positive) reach errors when fixating to the left of the target and leftward (negative) reach

errors when fixating to the right of the target. A similar pattern was observed in VIS/feedback trials, but with smaller magnitude (**Fig. 2b**; see also ref. 8). In contrast, during PROP trials, the reach endpoint was closer to the fixation point; for example, leftward errors were made when fixating to the left of the target (**Fig. 2e,f**). The gaze-dependent error patterns in the VIS + PROP trial types appeared to be a combination of those observed in VIS and PROP trials (**Fig. 2c,d**). For all trial types, the errors for the two gaze-locations aligned qualitatively when angular error was plotted in a retinotopic reference frame: that is, as a function of target relative to gaze (**Fig. 2**). This alignment might appear to support a retinotopic representation for reach planning^{8–10}. However, as these patterns differ markedly across trial types, they cannot be readily explained in terms of a fixed retinotopic bias, suggesting the need for a different explanation of these error patterns.

A planning model: integration across reference frames

We developed a model of reach planning that accounts for the pattern of gaze-dependent errors observed in our data. The model has two important 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, initial hand position and gaze location. As sensory signals are inherently variable, we modeled them as Gaussian likelihoods of true location given the sensory input, with likelihood variance reflecting the reliability of the sensory modality²⁸. Visual signals arrive in a retinotopic representation and proprioceptive signals arrive in a body-centered representation. Each available signal is then transformed into

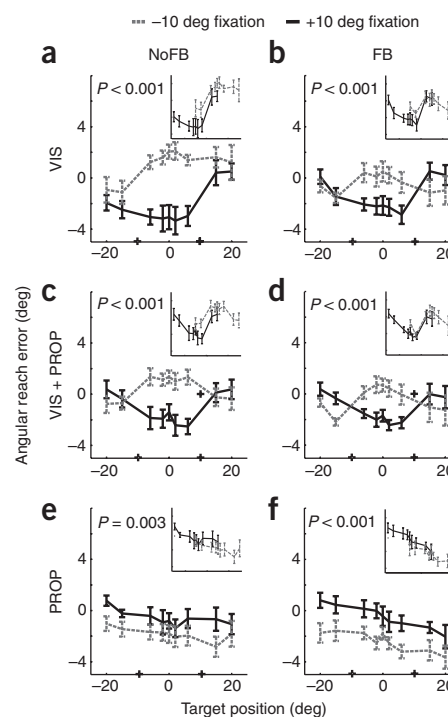


Figure 2 Average angular reach error across subjects for each trial condition. (a–f) Negative values indicate reach endpoints to the left of the target, and positive values indicate reach endpoints to the right of the target. Before averaging, linearly interpolated gaze-free errors for each trial type were subtracted (**Supplementary Section 1** and **Supplementary Fig. 4**). Error bars indicate standard errors. P values were determined with a paired permutation test⁵⁰. Insets show angular reach error plotted as a function of target position relative to gaze (that is, in a retinotopic coordinate frame).

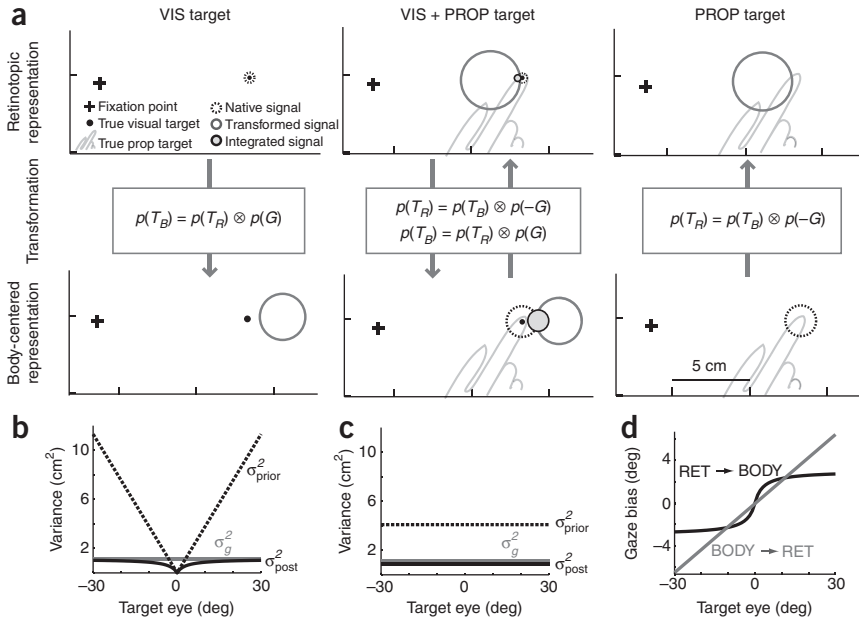


Figure 3 The bias and variance injected by sensory transformations in the model. **(a)** Examples of the bias and variance that arise during transformation of target information. Circles represent the posterior distribution of target location (95% confidence limits). Transformed signals have greater uncertainty and are biased either toward (retinotopic) or away from (body centered) the gaze location. In VIS + PROP trials, the native and transformed signals are integrated into a combined estimate of location (filled circles). **(b,c)** Posterior variance in estimated gaze location (transformation variance, black), gaze likelihood variance (gray) and variance of gaze prior (dashed) (see Online Methods) for target transformation from retinotopic to body-centered space **(b)** and target transformation from body-centered to retinotopic space **(c)**. **(d)** Bias in gaze estimate used in target transformations. Values in all panels were computed from INTEG model fit.

the non-native reference frame (**Fig. 3a**). Because subjects' head positions were fixed during the experiment, this complex nonlinear transformation²⁹ can be approximated simply by adding or subtracting the gaze location: that is, by convolving their distributions (**Fig. 3a** and Online Methods)¹. However, because the internal estimate of gaze location is also uncertain, this transformation adds variability to the signals. When both sensory modalities are available, the native and transformed signals are integrated in both reference frame representations (VIS + PROP target condition; **Fig. 3a**). Movement vectors are then computed in each representation by convolving the target and initial hand distributions (subtraction). It should be noted that, because the sensory transformation adds variability, each spatial variable is more reliably represented in one or the other of these representations depending on the availability and reliability of the relevant visual and proprioceptive signals (**Supplementary Section 2**).

The sensory transformations in our model also introduce a bias into the estimates of transformed variables. In particular, we posit that the internal estimate of gaze direction used to transform target location is biased toward the target. The bias was modeled as a Bayesian prior on gaze location¹ centered on the target. Because the transformation consists of adding or subtracting gaze location, depending on the direction of the transformation, the prior effectively biases the transformed target estimate either away from or toward the direction of gaze (**Fig. 3a**). Native (untransformed) target representations remain unbiased. Because the gaze prior is centered on the estimated target location, the variance of the prior (**Fig. 3b,c**) was assumed to scale with the uncertainty of the internal target estimate leading to the bias patterns shown in **Figure 3d**. Therefore, the gaze-dependent errors in the model depend on the relative weighting of the various sensory signals, depending on both the availability and reliability of sensory inputs and on the method of reading out the final movement vector.

We considered three possible output schemes for reading the planned movement vector from the model (**Supplementary Fig. 6**). Either 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 (INTEG) readout. The contributions of the two representations to the INTEG readout

depend on their relative reliability. It should be noted that these calculations use the simplifying assumption that the signals being combined are independent (**Supplementary Section 2**).

Each of the three potential readouts provided a quantitative prediction of reach errors as a function of hand, target and gaze locations. The only free parameters in the model were the variances of individual sensory inputs and the gaze priors (Online Methods). The values of the proprioceptive variances were based on previously reported values³⁰. Four parameters remained: 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

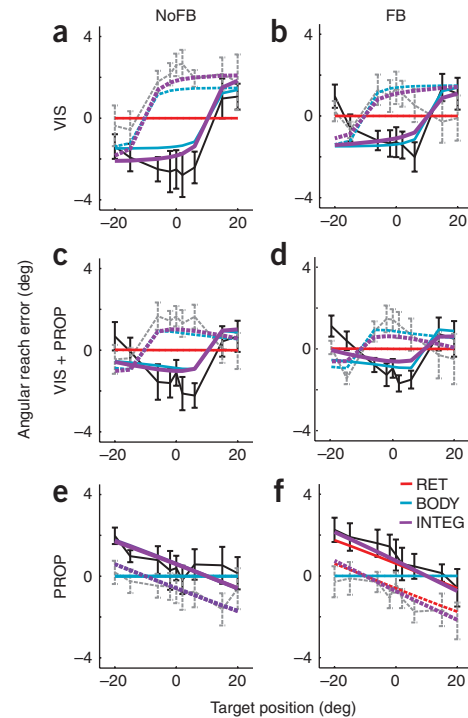


Figure 4 Model fits for gaze-dependent reach errors. **(a-f)** Black and gray lines show mean (standard error) errors across subjects for each trial condition, after subtracting the overall mean separately for each of the six trial types. Colored lines show best-fit model predictions. Solid lines, gaze right; dashed lines, gaze left.

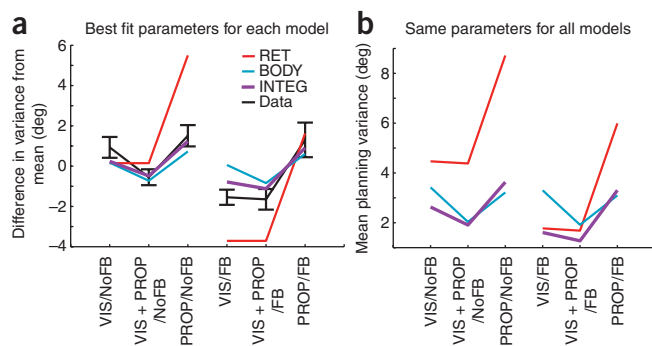


Figure 5 Reach variability. **(a)** Trial-type differences in angular reach variance. Black lines represent mean (standard error) reach variance across subjects. Values are derived from the average variance across subjects and trial conditions in each trial type, with the overall mean subtracted. Colored lines represent model predictions with each readouts fit parameters. **(b)** Mean variability of each model readout scheme, as a function of trial type, using fixed model parameters (INTEG fit).

the gaze-dependent errors (Fig. 2) after mean correction (see Online Methods) using least-squares regression (Supplementary Section 2, Supplementary Table 1 and Supplementary Fig. 7).

Model fits of constant and variable reach errors

We first considered how well the three output models fit the observed patterns of gaze-dependent errors (Fig. 4a–f). When the model was fit with a single-representation readout, RET or BODY, it failed to predict errors for all trial types. This is because only transformed target signals contain a gaze-dependent error, and thus when the readout was in the native reference frame of a unimodal target, no error was predicted (Fig. 4a,b,e,f). Although both readouts had errors for bimodal targets (VIS + PROP; Fig. 4c,d), errors in the RET readout were a result of the transformed proprioceptive signal and would therefore be in the wrong direction (Fig. 3a). Thus, for the RET readout, the fitting procedure nullified the effect of the transformed proprioceptive signal by forcing the model to rely only on vision: that is, by driving the visual variance parameter toward zero (Supplementary Table 1). In contrast, when the movement vector representations were combined in the INTEG readout, the model performed well across all trial types. This readout captured both changes in the magnitude of gaze-dependent errors (VIS versus VIS+PROP and feedback versus NoFB) and the sign reversal that we observed with PROP targets. It accomplished this by differentially weighting the two reference frames across trial conditions, with both representations making substantive contributions (Supplementary Section 2 and Supplementary Fig. 8). Indeed, a scheme that simply switches between the RET and BODY readouts depending on task would not capture the data well. First, both the RET and BODY readouts failed to capture the differences in the magnitude of gaze-dependent errors that we observed between feedback and NoFB conditions. Second, to predict the observed errors, the switching scheme would need to rely predominantly on the more variable transformed signals, rather than on native signals, a sub-optimal arrangement.

In addition to fitting the gaze-dependent error patterns, our model predicted the

differences in movement variability across trial types (Fig. 5a). Because computations in the model were assumed to be noise free, model output variability was entirely the result of variability in the sensory inputs, shaped by the model computations, and did not require any additional parameter fitting. The INTEG model fit provided an accurate prediction for the changes in output variance across trial types and was better than the two single-representation fits. These predictions came from separate parameter fits for each readout. However, the model parameters presumably reflect actual variances in the neuronal representations of sensory inputs. Using a single set of variances, for example, the INTEG fit, we looked at how variability in the movement plan depends on readout. We found that the INTEG readout generally yielded a lower variance estimate (Fig. 5b), as it made better use of all available sensory signals (although the extent of this advantage depended on the statistical properties of the sensory transformations; Supplementary Section 2 and Supplementary Fig. 9). In contrast to the idea that a single coordinate frame should dominate movement planning^{3–10,15,18}, this analysis illustrates that using multiple representations of a movement plan yields more reliable performance across tasks.

Model predictions for previously published datasets

We tested the model, fit to our own dataset, on a similar, previously published dataset⁸ containing visual target trials with an expanded range of movements (that is, more start and gaze locations). These data showed the retinal eccentricity effect, and, as described above, the effect magnitude was smaller when visual feedback of the hand was available (data reproduced in Fig. 6a,b). In addition, there was a component of the reach error that correlated with the relative positions of the hand and target (data reproduced Fig. 6c,d). Our model captured all of these important features in the dataset (Fig. 6e–h).

Our model explains another very different empirical result, again without additional parameter fitting. We have previously shown that visual information about initial hand location is weighted more heavily when reaching to visual targets (as in VIS/feedback trials) than when reaching to proprioceptive targets (as in PROP/feedback trials)². We previously proposed that this sensory reweighting was a result of the cost (for example, variability) incurred by sensory transformations². The present model makes this cost explicit, with quantitative predictions of the angular error that should result from artificial shifts in visual feedback. In both the empirical data and the INTEG readout predictions, visual feedback shifts had a weaker effect when reaching to proprioceptive targets than when reaching to visual targets (Fig. 7a). This effect was

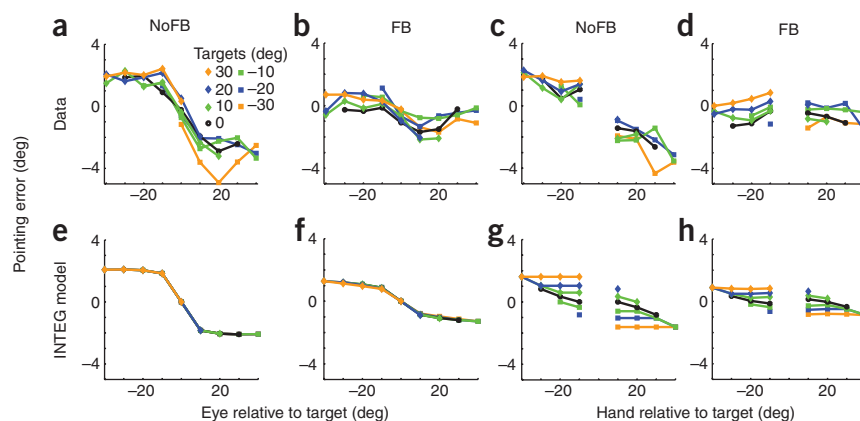


Figure 6 Model predictions for data from a previous study⁸ (a,b,e,f). Gaze-dependent pointing error. (c,d,g,h) Pointing error as a function of initial hand position relative to the target. The average pointing error from a previous study⁸ is shown in a–d. Our INTEG model predictions of pointing error with parameters fit to our data are shown in e–h.

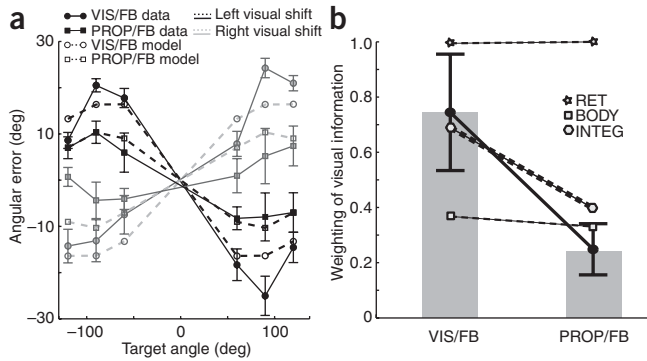


Figure 7 Changes in sensory weighting with target modality. **(a)** Mean angular error induced by artificial shifts in the visual feedback of the hand before movement onset. Data are from a previous study². Error bars represent standard error. Model predictions use the INTEG readout, with parameters fit to our data. **(b)** Relative weighting of visual versus proprioceptive information about initial hand position in movement planning for reaches to VIS and PROP targets. Error bars represent s.d. across subjects. Dashed lines show model predictions for each readout scheme.

quantified in terms of overall weighting of visual versus proprioceptive feedback (**Fig. 7b**), which was much greater for VIS targets than for PROP targets. In the model, this re-weighting was a result of the tradeoff between the retinotopic and body-centered representations of the movement plan (**Supplementary Fig. 8**), 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.

Origins of the gaze bias

We found that a bias in the internal estimate of gaze location can account for the complex pattern of gaze-dependent reach errors that we observed across trial types. This bias might arise as a result of either 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 (**Supplementary Section 3** and **Supplementary Fig. 10**). 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 linked³¹. Indeed, the model bias was implemented in this manner (see Online Methods).

Although the source of the gaze bias remains undetermined, we were able to corroborate its presence using an independent measure. Specifically, we found that a visually peripheral reach target biased a subject’s estimate of straight ahead, and this bias was consistent with a shift in the estimated gaze direction³² toward the target (**Supplementary Section 3** and **Supplementary Fig. 11**). This perceptual effect, similar to the errors observed in our reach experiment, was well modeled by a Bayesian prior on gaze location centered at the target.

DISCUSSION

Our study sought to test two widely held ideas in the field of sensorimotor control: that spatial patterns of errors for a given movement reflect the underlying neural representation^{7–13} and that a single reference frame should dominate movement planning^{3–6,8–10,15,18,33}. We argue that neither of these ideas is correct. First, we found that a single, apparently retinotopic, pattern of reach errors could be explained by a model in which multiple neural representations are used (for example, a combination of both retinotopic and body-centered reference frames). Second, we found that using more than one representation conferred an advantage in terms of reduced planning variability.

Spatial patterns of reach errors, especially retinotopic or gaze-centered patterns, have been cited as evidence that the neural representation for reach planning occurs in a particular reference frame^{5,8–13,33}. We observed similar error patterns, but found that their magnitude and directionality varied with the sensory signals that were available for movement planning. This variation was inconsistent with a fixed bias in a single neural representation. We found that a bias in the transformation of target information between representations could lead to gaze-dependent error patterns in both retinotopic and body-centered representations. In fact, the retinal eccentricity effect for visual targets arose in the model only from the body-centered representation, clearly illustrating that the spatial pattern of reach errors need not be a good indicator of the underlying neural representation. More generally, we argue that when sensory signals are used in a statistically optimal manner^{28,34–38}, the same information is contained in multiple neural representations and there is no required relationship between the behavioral output and any single representation of the movement plan.

We propose that a biased transformation can account for gaze-dependent reaching errors, focusing on azimuthal (left-right) reach errors and azimuthal gaze shifts. When gaze is shifted vertically (up-down), a similar pattern of vertical gaze-dependent errors is observed³⁹, which is qualitatively consistent with our model. We did not observe gaze-dependent effects in the radial (depth) reach errors of our subjects (**Supplementary Fig. 4**), probably because gaze was not varied in depth. When this manipulation is done, a complex pattern of gaze-dependent depth errors emerges³³, presumably reflecting the complex binocular, three-dimensional geometry of the eyes²⁹. Modeling these patterns will require substantially more complex model representations and transformations.

Although we have focused on gaze-dependent errors, a variety of other error patterns are amenable to similar analysis. For example, several studies exploring the effect of head rotations on reach accuracy have found errors in reaching toward the direction in which the head is oriented^{40,41}, a pattern that is consistent with a bias of the perceived midline of the head toward the direction of gaze⁴¹. In addition, the leftward bias in reaching that we observed (**Fig. 2**) can be modeled as a rightward bias in the proprioceptive estimate of the right hand (**Supplementary Section 1**). The idiosyncratic gaze-independent error patterns exhibited by individual subjects (**Supplementary Fig. 2**) may also be explained by subject-specific biases on sensory variables⁴². All of these biases could be the result of prior expectations on sensory variables or their correlations³⁶, although alternate sources of errors are plausible, such as impoverished neural representations⁴³.

We argue that the variability of sensory transformations make it advantageous to use multiple representations for a movement plan. First, we found that using multiple reference frames in a weighted fashion improved movement variability across trial types (**Fig. 5b**). Next, we found that only multiple reference frames in the model could account for the reweighting of visual feedback of the hand as a result of target type (**Fig. 7**)². This is because the relative variability of the two representations, and hence their contribution to the output, depends on the sensory modality of the target. However, it should be noted that the model does not account for all of the observed reweighting (**Fig. 7**). This may be explained by a reduced weighting of vision as a result of a lack of belief by the subject that the visual cursor is at their finger⁴⁴. Still, our results suggest that, although a single representation may simplify the flow of information^{3–6}, it does not make optimal use of that information in movement planning.

Previous studies have reported patterns of movement errors^{2,12,26} or generalization of motor learning^{45,46} that cannot be explained succinctly in a single reference frame. Such error patterns often depend on the

availability of sensory signals^{2,12} and are explained in terms of changes in the underlying reference frame^{44,46} or an intermediate representation of movement planning^{26,45}. Our results provide a different explanation; movements are always represented in multiple reference frames, independent of the task, and it is the statistical reliability of these representations that 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^{24,47}. Reach-related areas have been found to have retinotopic coding^{15,48}, head- or body-centered coding^{19,49}, and hand- and shoulder-centered coding^{16,21–23}, as well as mixed representations^{14,16–18,21–23,49}. The two separate representations of movement plan in our model could be found in subsets of these cortical areas. Alternatively, the same computation could be performed using a heterogeneous neural population that contains both retinotopic and body-centered components, that is, mixed representation. Indeed, these implementations are at two ends of a continuum and physiological evidence seems to point to a middle ground in which all of the cortical areas for reach planning make use of mixed representations, but the parietal cortex has a more retinotopic character and the frontal cortex is more hand or body centered^{24,47}. As our results indicate, 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^{4,14,49}.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/natureneuroscience/>.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We would like to thank S.M. Beurze, S. van Pelt, W.P. Medendorp and S.J. Sober for generously providing their data for model comparison. 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.

Published online at <http://www.nature.com/natureneuroscience/>.

Reprints and permissions information is available online at <http://www.nature.com/reprintsandpermissions/>.

- Schlicht, E.J. & Schrater, P.R. Impact of coordinate transformation uncertainty on human sensorimotor control. *J. Neurophysiol.* **97**, 4203–4214 (2007).
- Sober, S.J. & Sabes, P.N. Flexible strategies for sensory integration during motor planning. *Nat. Neurosci.* **8**, 490–497 (2005).
- 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).
- 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).
- Engel, K.C., Flanders, M. & Soechting, J.F. Oculocentric frames of reference for limb movement. *Arch. Ital. Biol.* **140**, 211–219 (2002).
- Lacquaniti, F. & Caminiti, R. Visuo-motor transformations for arm reaching. *Eur. J. Neurosci.* **10**, 195–203 (1998).
- Soechting, J.F. & Flanders, M. Sensorimotor representations for pointing to targets in three-dimensional space. *J. Neurophysiol.* **62**, 582–594 (1989).
- Beurze, S.M., Van Pelt, S. & Medendorp, W.P. Behavioral reference frames for planning human reaching movements. *J. Neurophysiol.* **96**, 352–362 (2006).
- 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).
- Pouget, A., Ducom, J.C., Torri, J. & Bavelier, D. Multisensory spatial representations in eye-centered coordinates for reaching. *Cognition* **83**, B1–B11 (2002).
- 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).
- 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).
- 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).
- 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).
- Batista, A.P., Buneo, C.A., Snyder, L.H. & Andersen, R.A. Reach plans in eye-centered coordinates. *Science* **285**, 257–260 (1999).
- Batista, A.P. *et al.* Reference frames for reach planning in macaque dorsal premotor cortex. *J. Neurophysiol.* **98**, 966–983 (2007).
- 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).
- Buneo, C.A., Jarvis, M.R., Batista, A.P. & Andersen, R.A. Direct visuomotor transformations for reaching. *Nature* **416**, 632–636 (2002).
- 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).
- 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).
- 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).
- Wu, W. & Hatsopoulos, N. Evidence against a single coordinate system representation in the motor cortex. *Exp. Brain Res.* **175**, 197–210 (2006).
- Wu, W. & Hatsopoulos, N.G. Coordinate system representations of movement direction in the premotor cortex. *Exp. Brain Res.* **176**, 652–657 (2007).
- Burnod, Y. *et al.* Parieto-frontal coding of reaching: an integrated framework. *Exp. Brain Res.* **129**, 325–346 (1999).
- Caminiti, R., Ferraina, S. & Mayer, A.B. Visuomotor transformations: early cortical mechanisms of reaching. *Curr. Opin. Neurobiol.* **8**, 753–761 (1998).
- Carrozzo, M. & Lacquaniti, F. A hybrid frame of reference for visuo-manual coordination. *Neuroreport* **5**, 453–456 (1994).
- Bock, O. Contribution of retinal versus extraretinal signals towards visual localization in goal-directed movements. *Exp. Brain Res.* **64**, 476–482 (1986).
- Knill, D.C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
- Blohm, G. & Crawford, J.D. Computations for geometrically accurate visually guided reaching in 3-D space. *J. Vis.* **7**, 1–22 (2007).
- 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).
- Land, M.F. & Hayhoe, M. In what ways do eye movements contribute to everyday activities? *Vision Res.* **41**, 3559–3565 (2001).
- Balslev, D. & Miall, R.C. Eye position representation in human anterior parietal cortex. *J. Neurosci.* **28**, 8968–8972 (2008).
- Van Pelt, S. & Medendorp, W.P. Updating target distance across eye movements in depth. *J. Neurophysiol.* **99**, 2281–2290 (2008).
- Ernst, M.O. & Banks, M.S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
- Helbig, H.B. & Ernst, M.O. Optimal integration of shape information from vision and touch. *Exp. Brain Res.* **179**, 595–606 (2007).
- Kording, K.P. & Wolpert, D.M. Probabilistic mechanisms in sensorimotor control. *Novartis Found. Symp.* **270**, 191–198; discussion 198–202, 232–197 (2006).
- 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).
- Yuille, A.L. & Bulthoff, H.H. Bayesian decision theory and psychophysics. in *Perception as Bayesian Inference* 123–136 (Cambridge University Press, Cambridge, UK, 1996).
- 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).
- 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).
- Lewald, J. & Ehrenstein, W.H. Visual and proprioceptive shifts in perceived egocentric direction induced by eye position. *Vision Res.* **40**, 539–547 (2000).
- 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).
- Soechting, J.F. & Flanders, M. Errors in pointing are due to approximations in sensorimotor transformations. *J. Neurophysiol.* **62**, 595–608 (1989).
- Kording, K.P. & Tenenbaum, J.B. Causal inference in multisensory integration. in *NIPS 737–744* (MIT Press, Cambridge, Massachusetts, 2006).
- Ahmed, A.A., Wolpert, D.M. & Flanagan, J.R. Flexible representations of dynamics are used in object manipulation. *Curr. Biol.* **18**, 763–768 (2008).
- 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).
- 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).
- Cohen, Y.E. & Andersen, R.A. Reaches to sounds encoded in an eye-centered reference frame. *Neuron* **27**, 647–652 (2000).
- 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).
- Good, P.I. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*, 2nd edn. (Springer-Verlag, New York, 2000).

ONLINE METHODS

Experimental setup. Subjects were seated in a virtual reality setup (Supplementary Fig. 1). The right arm rested on top 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 of the table with the wrist supine. Thus, the two hands never came in contact. The location of both index fingers was monitored using an infrared tracker (Optotrak 3020, Northern Digital). The subjects' view of their arms was blocked by a mirror through which they viewed a rear-projection screen (Supplementary Fig. 1). This provided the illusion that visual objects 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 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 (Supplementary Fig. 1). The targets (8-mm radius green disks, when visual) were located at $\pm 20^\circ$, $\pm 15^\circ$, $\pm 6^\circ$, $\pm 2^\circ$ and 0° relative to midline. Two gaze fixation points (5-mm radius red disks) were located on this arc at $\pm 10^\circ$. Visual feedback was given with an 8-mm radius disk centered on the index finger: white for the right hand and blue for the left.

All trials consisted of four steps. First, subjects moved their right index finger to a fixed starting location. On feedback trials, the start location was indicated with a 10-mm radius green disk and visual feedback of the right hand was illuminated. On NoFB trials, neither feedback nor target were visible and subjects were guided to the start location using the arrow field method, which provides no feedback of absolute hand position². Second, when the right index finger came to rest within 10 mm of the start location, one of the fixation points appeared. Subjects were required to maintain fixation at the fixation point for the remainder of the trial. Third, after fixation, 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 and subjects moved the left index finger to the target, and the target disk was then extinguished leaving the blue feedback disk. For PROP trials, an arrow field was used to guide the unseen left hand to the unseen target location. Fourth, after target specification, there was a 500-ms delay before an audible go tone was played and subjects reached to the target. On feedback trials, both feedback and the start disk were extinguished at 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 females and six males) participated in the experiment. Subjects were right-handed, had no known neural or motor deficits, had normal or corrected-to-normal vision, and gave written informed consent before participation. The experiment was divided into two sessions, which were performed on different days to minimize fatigue. One session contained feedback trials and the other contained NoFB trials, and session order was randomized across subjects. Each session contained six repetitions of each of the 54 conditions (three target types \times nine targets \times two fixation points), for a total of 324 trials (not including error trials, which were repeated). These trials were followed by trials in which gaze location was unconstrained. This set consisted of six repetitions of nine trial conditions (three target types \times three targets), bringing the total number of trials for each session to 378. The order of presentation across conditions was randomized in each repetition.

Data analysis. For each trial, reach endpoint was defined as the position at which movement speed first fell to 5 mm s^{-1} . Reach targets and endpoints were converted into polar coordinates about an origin located directly below the midpoint of the two eyes. Angular reach error is the angular difference between the endpoint and target, with positive values indicating 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, the angular reach errors were corrected by subtracting the linearly interpolated free-gaze errors (separately for each subject and trial type) to minimize the effects of idiosyncratic gaze-independent errors while preserving the relationship between error, gaze and target (see Supplementary Figs. 3–5). The significance of gaze-dependent effects was tested by a paired permutation test of main effect of gaze location⁵⁰.

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 , that is, Gaussian likelihoods, $N(X, \sigma^2)$, with an isotropic covariance matrix of $\sigma^2 I$: vision of the right fingertip is represented by $p(f_v | F_R) \sim N(F_R, \sigma_v^2)$, proprioception of the right fingertip is represented by $p(f_p | F_B) \sim N(F_B, \sigma_p^2)$, vision of the target is represented by $p(t_v | T_R) \sim N(T_R, \sigma_v^2)$, proprioception of the target is represented by $p(t_p | T_B) \sim N(T_B, \sigma_p^2)$ and the felt gaze position is represented by $p(g | G) \sim N(G, \sigma_g^2)$ (equation (1)). Lower-case variables are sensory signals, with subscripts denoting sensory modality (v for visual and p for proprioceptive). Upper-case variables are true locations with subscripts denoting reference frame (R for the retinotopic location and B for body-centered location).

When a signal is unavailable in a given trial type (for example, f_v in NoFB trials), the likelihood is 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 uncertainty in X given x , that is, on the posterior distributions $p(X|x)$. Bayes' rule relates these two distributions:

$$p(X|x) \propto p(x|X)p(X) \quad (2)$$

where $p(X)$ represents prior information about the location. We usually assume that the prior is flat, so the posterior is proportional to the likelihood (although a nontrivial prior is described below).

All of the computations in the model make locally optimal use of the signals required for a computation, assuming that those signals are independent. Indeed, only two operations are performed by the network: signal integration and addition (or subtraction). Integration is the process of combining 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 the product of the two input distributions:

$$p(X|x_1, x_2) \propto p(X|x_1)p(X|x_2) \quad (3)$$

The resulting posterior is also Gaussian, with mean and variance:

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

In equation (4) and below, $\mu_{X|x}$ and $\sigma_{X|x}^2$ represent the mean and variance of the distribution $p(X|x)$, respectively. If either input distribution is uniform, that is, 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. 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 convolving the inputs:

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

The result is also a Gaussian:

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

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 fingertip and target locations in both retinotopic and body-centered reference frames. These representations integrate all available sensory signals, requiring the transformation of non-native signals. For example, in computing the retinotopic representation of target, the proprioceptive signal t_p must be transformed. Because head position is fixed, $T_R = T_B - G$, and the transformation follows equation (5):

$$p(T_R | t_p, g) = p(T_B | t_p) \otimes p(-G | g) \quad (7)$$

Parallel transformations 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:

$$p(T_R | t_v, t_p, g) \propto p(T_R | t_v) p(T_R | t_p, g) \quad (8)$$

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

The model next computes retinotopic and body-centered representations of the desired movement vector. Because the true value of the instructed movement vector is $MV = T - F$, these computations also follow equation (5):

$$p(MV_U | t_v, t_p, f_v, f_p, g) = p(T_U | t_v, t_p, g) \otimes p(F_U | f_v, f_p, g), \quad (9)$$

$$U = R, B$$

Note that both of the inputs to equation (9) depend on gaze. If the same estimate of gaze is used in all transformations, then the inputs are not fully independent, and equation (6) is only an approximate solution (see **Supplementary Section 2** and **Supplementary Fig. 9**).

Finally, 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:

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

Note that these three readouts are maximum a posteriori estimates given the respective posteriors. This computation is only correct when the input signals are independent, which is not always true (see **Supplementary Section 2**).

The final component of the model is a bias in the transformation of target position between reference frames. We modeled this bias as a systematic mis-estimation of gaze location as a result of a Bayesian prior. The prior takes the form of a Gaussian distribution, $p(G) \sim N(t_{\text{prior}}, \sigma^2_{\text{prior}})$. The mean of the prior distribution t_{prior} is a Gaussian random variable with mean T_B and variance proportional to that of the transformed target variable. The variance of the prior, σ^2_{prior} is a product of the variance of the target distribution being transformed

and a scaling factor that depends on the modality of the transformed signal (see **Supplementary Table 1**). This prior effects the gaze estimate used for transforming target locations, but not the transformation of finger locations.

Because the model is composed entirely of integration (equation (3)) and addition (equation (5)) operations, all expected values, including the movement vector readouts, can be written as weighted sums of the means of the initial sensory inputs, with coefficients depending only on the variances of those signals (equations (4) and (6)). The trial-by-trial variances of the readouts (**Fig. 6**) can be computed from these coefficients and the input variances in equation (1).

Model fitting. The only model parameters are the sensory variances in equation (1) and the variance of the gaze prior, σ^2_{prior} . The proprioceptive variances were set a priori on the basis of previously published estimates³⁰. The variance of visual signals, σ^2_v , is assumed to scale linearly with the distance of the stimulus from center of gaze and this scale factor is the first free parameter. The variance of the gaze signal, σ^2_g is the second. 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 (**Fig. 4**). The model generates Cartesian movement vectors, which are then converted into polar coordinates to obtain angular errors. The fitting procedure minimized the sum square prediction error across trial conditions using the Matlab optimization toolbox (function `fmincon`, Mathworks). Optimization was repeated 100 times with random initial parameter values, and the final parameters were largely insensitive to initial values.

Modeling other datasets. The model predictions for previously published datasets (**Figs. 6** and **7**) used the model parameters fit to our own dataset with the INTEG readout. When modeling the reach errors for a previous study⁸ (**Fig. 6**), we translated their target array into our workspace and rescaled the spacing of the targets to maintain the same azimuthal separation. The data for **Figure 7** come from the gaze-fixed trials in another previous study² (**Supplementary Fig. 3**).