A Gain-Field Encoding of Limb Position and Velocity in the Internal Model of Arm Dynamics

Eun Jung Hwang*, Opher Donchin, Maurice A. Smith, Reza Shadmehr

Laboratory for Computational Motor Control, Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland, United States of America

Adaptability of reaching movements depends on a computation in the brain that transforms sensory cues, such as those that indicate the position and velocity of the arm, into motor commands. Theoretical consideration shows that the encoding properties of neural elements implementing this transformation dictate how errors should generalize from one limb position and velocity to another. To estimate how sensory cues are encoded by these neural elements, we designed experiments that quantified spatial generalization in environments where forces depended on both position and velocity of the limb. The patterns of error generalization suggest that the neural elements that compute the transformation encode limb position and velocity in intrinsic coordinates via a gain-field; i.e., the elements have directionally dependent tuning that is modulated monotonically with limb position. The gain-field encoding makes the counterintuitive prediction of hypergeneralization: there should be growing extrapolation beyond the trained workspace. Furthermore, nonmonotonic force patterns should be more difficult to learn than monotonic ones. We confirmed these predictions experimentally.

Introduction

Behavioral (Shadmehr and Mussa-Ivaldi 1994; Conditt and Mussa-Ivaldi 1999) and neurophysiological (Li et al. 2001; Gribble and Scott 2002) evidence suggests that the brain controls reaching movements with highly adaptable internal models that predict behavior of the limb as it interacts with forces in the external world. To infer how the brain learns internal models, research has been conducted in three fields: psychophysics, neurophysiology, and computational modeling. Psychophysical experiments have quantified generalization, i.e., how error experienced in one movement (in a given position and direction) affects neighboring movements. It appears that there is a specific pattern to how the brain generalizes movement errors to other directions (Sainburg et al. 1999; Thoroughman and Shadmehr 2000), to other arm configurations (Shadmehr and Moussavi 2000; Malfait et al. 2002), and to movements with different trajectories (Conditt et al. 1997; Goodbody and Wolpert 1998).

Neurophysiological experiments have suggested that the motor cortex may be one of the crucial components of the neural system that learns internal models of limb dynamics (Li et al. 2001). There is now significant information about how various movement parameters, e.g., limb velocity (Moran and Schwartz 1999), arm orientation (Scott and Kalaska 1997; Scott et al. 1997), and hand position (Caminiti et al. 1990; Sergio and Kalaska 1997) are encoded by neurons in the motor cortex.

Computational models with elements reflecting some of the cell properties found in neurophysiological experiments have attempted to explain how patterns of generalization during adaptation may be related to the neural representation. These computational models hypothesize that an internal model is composed of "elements," or bases, each encoding only part of sensory space, and that population codes combine these elements when computing sensorimotor transformations (Georgopoulos et al. 1986; Levi and Camhi

2000; Pouget et al. 2000; Thoroughman and Shadmehr 2000; Donchin and Shadmehr 2002; Steinberg et al. 2002). This map would transform a desired sensory state into a prediction of upcoming force. Under these assumptions, patterns of error generalization should reveal the shape of the basis elements.

We performed a set of experiments that examined how the neural elements might simultaneously encode limb position and velocity. We show that movement errors generalize with a pattern that suggests a linear or monotonic encoding of limb position space and that this encoding is multiplicatively modulated by an encoding of movement direction. The gain-field encoding of limb position and velocity that we infer from the generalization patterns is strikingly similar to neural encoding of these parameters in the motor cortex (Georgopoulos et al. 1984).

Results

Adapting to a Position- and Velocity-Dependent Field

Figure 1 describes an experiment in which subjects performed reaching movements in force fields that depended on both velocity and position of the limb. Subjects made movements in the horizontal plane while holding the handle of a robot. The task was to reach a target (displacement of approximately 10 cm; see Materials and Methods) within 500 ± 50 ms. Handle and target positions were continuously

Abbreviations: d.f., degree of freedom; PE, perpendicular error; SEM, standard error of the mean

Academic Editor: James Ashe, University of Minnesota

* To whom correspondence should be addressed. E-mail: ehwang@bme.jhu.edu
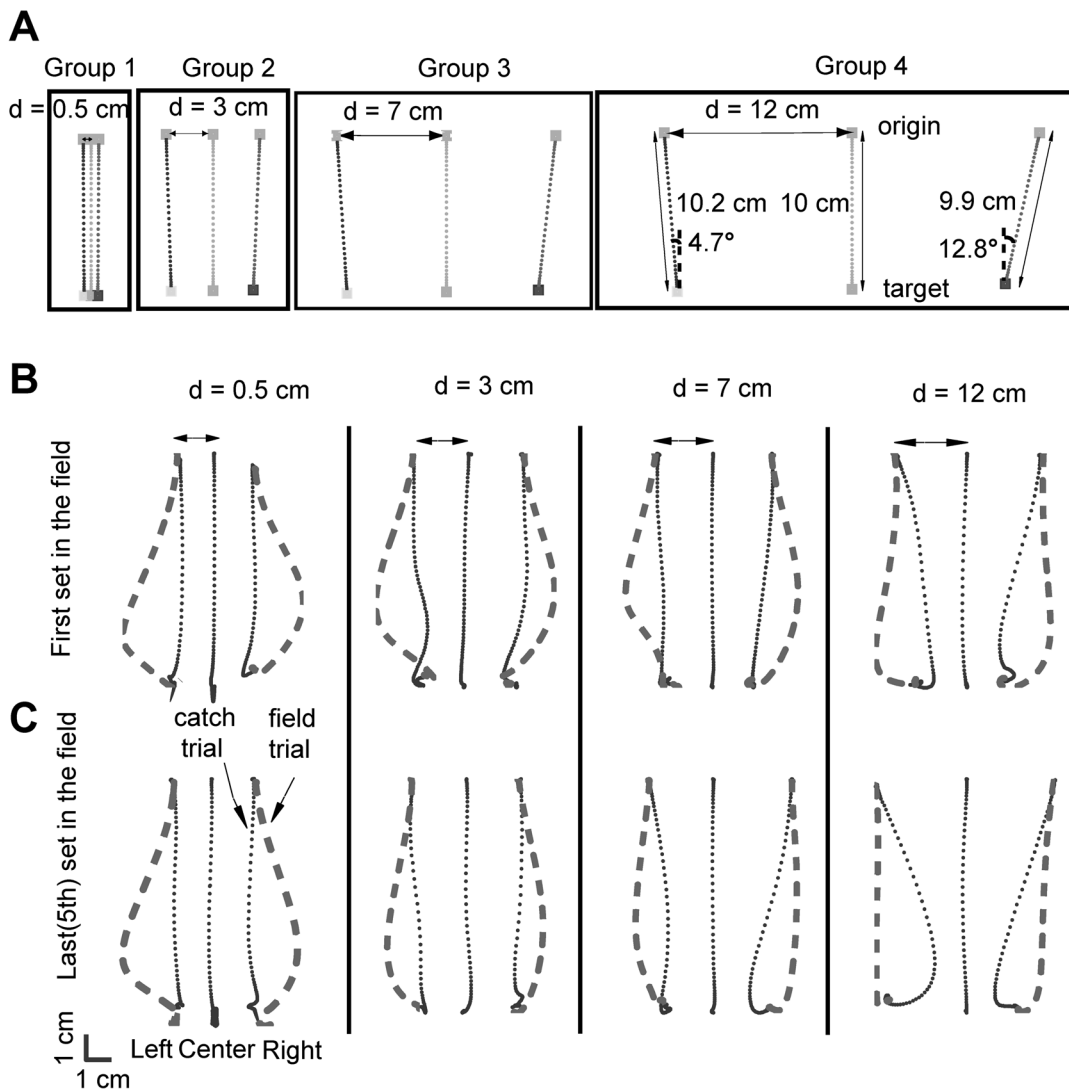
Figure 1. Adaptation to a Force Field That Depends on Both Position and Velocity of the Limb

(A) The origin of the center movements is aligned with the subject's body midline, and the origins of the left and right movements are symmetrically positioned with a given separation distance (d) for each group. The target positions shown here are an example for a subject with typical arm lengths (20, 33, 34 cm shoulder, upper arm, and lower arm lengths, respectively). In order to help subjects distinguish among locations, different colors were provided for targets at different locations (yellow, green, and blue for the left, center, and right, respectively).
(B) The average trajectories in three positions—left, center, and right (one subject per column)—for the first third of the movements from the first field set (trials 1–28). Dashed lines are movements during which force field is on and dotted lines are catch trials. Separation distances between neighboring movements (d) are not scaled in this figure.
(C) The average trajectories for the first third of the fifth field set (trials 337–364). The task is much easier to learn when the three movements are spatially separated from each other.
DOI: 10.1371/journal.pbio.0000025.g001

projected onto a screen placed directly above the subject's hand (Shadmehr and Moussavi 2000). Feedback on performance was provided immediately after target acquisition, but feedback related only to the subject's success in arriving at the target within the prescribed time window and not to the shape of the hand trajectories. After completion of each movement, the robot moved the hand to a new start position and another target was presented. The start positions were pseudorandomly chosen from three possible locations: left, center, and right. Twenty-four subjects were divided into four groups, and the start positions for the four groups were separated by 0.5 cm, 3 cm, 7 cm, or 12 cm, respectively (Figure

1A). Different colors were used for the left, center, and right start positions and targets so that we could be certain that subjects could distinguish the locations even when the separation distances were small. The targets were placed so that movements from all three starting locations required the same joint angle displacement. Thus, the movements were parallel in joint space and not in Cartesian space (Figure 1A). Therefore, the movements explored the same joint velocity space but at different joint positions.

The robot could apply arbitrary patterns of force to the hand. We programmed it so that the movements were perturbed by a viscous curl-field. In a viscous curl-field, the

force is proportional to speed and perpendicular to velocity. However, our viscous curl-fields also depended on position. During movements from the left starting position, the robot perturbed the hand with a clockwise curl-field ($B = [0\ 13;\ -13\ 0]$ N·s/m, pushing the arm leftward during the movement). For movements starting on the right, a counterclockwise curl-field ($B = [0\ -13;\ 13\ 0]$ N·s/m) was present (pushing the arm rightward). For the center movements, the field was always null (no forces were applied). Thus, to succeed in the task, the subjects needed to produce three different force patterns although the movements required the same joint angular velocities. The idea was to find out how far apart the movements needed to be in position space for the task to become learnable. To familiarize the volunteers with the task and produce baseline performance, subjects first did three sets of 84 movements in which no forces were applied. Following the baseline sets, subjects did five force-field sets.

We found a limited ability to adapt to such position-dependent viscous curl-fields. Figure 1B and 1C displays the average hand paths of movements during the first and last sets of training for typical subjects from each group. The figures show movements both in field trials and in catch trials (occasional trials interspersed with the field trials in which the robot did not apply any forces). In the early phase of training, field trials were strongly curved toward the direction of force, although slight adaptation appeared in the largest separation distance group ($d = 12$ cm; Figure 1B). Late in training (Figure 1C), subjects with the largest separation in starting position showed manifest adaptation. Hand paths in the field trials became straighter, and trajectories of catch trials showed large aftereffects (Figure 1C). In contrast, subjects with the smallest separation in starting position showed little improvement in performance. Therefore, it appeared that movements that were spatially close to each other could not be easily associated with different force patterns. As the movements became farther apart, different forces could be more readily associated with them.

As a measure of error, we used displacement perpendicular to target direction at 250 ms into the movement (perpendicular error [PE]). Figure 2 shows the error on each trial averaged across subjects and plotted in a time series for Group 1 ($d = 0.5$ cm; Figure 2A) and Group 4 ($d = 12$ cm; Figure 2B). A gradual decrease in error magnitude and an increase in aftereffects in catch trials were apparent in Group 4, but not in Group 1. A learning index combining performance on field and catch trials (see Materials and Methods) allows a comparison of performance across groups (Figure 2C). An ANOVA on the learning index showed a significant effect both for separation distance and set number ($F = 41.78$, d.f. $= 3$, $p < 1.0 \times 10^{-8}$ for distance factor; $F = 3.02$, d.f. $= 4$, $p < 0.02$ for set number factor), suggesting that subjects performed better in the groups where targets were spatially separated.

We were struck by another difference between Figure 2A and 2B. There were never any forces for the center movement. However, the variance in these movements (middle traces) changed when going from the baseline sets to the field sets. For instance, the center movements in Group 1 have a much larger variability in field sets than in baseline sets. Our interpretation of this is that the forces subjects experienced on the left and the right influenced the center movement through generalization (see Dataset S1, found at

http://dx.doi.org/10.1371/journal.pbio.0000025.sd001, for a trial-by-trial analysis). We quantified this generalization (or interference) to center movements using an index (see Materials and Methods). In Figure 2D, the generalization index is shown for all groups. An ANOVA on separation distance by set number shows that generalization varies significantly with separation distance but not with set number ($F = 15.56$, d.f. $= 3$, $p < 2.2 \times 10^{-8}$ for the distance factor; $F = 0.83$, d.f. $= 4$, $p > 0.5$ for set number factor). As neighboring movements became spatially farther apart, generalization among them appeared to decrease.

Accounting for the Experimental Data with a Model

The above results demonstrate that when different forces are to be associated with two movements that are in the same direction but at different spatial locations, generalization decreases with increased distance between them. On the other hand, earlier results had found that when movements to various directions are learned at a single location, learning generalizes to other arm locations very far away (Ghez et al. 2000; Shadmehr and Moussavi 2000; Malfait et al. 2002).

To reconcile these two apparently contradictory findings, we performed a simulation of the internal model in which the force field was represented as a population code via a weighted sum of basis elements. Each element was sensitive to both the position and velocity of the arm. The crucial question was how each element should code limb position and velocity to best account for all the available data on generalization. Previous work had shown that velocity encoding was consistent with Gaussian-like functions (Thoroughman and Shadmehr 2000). To account for both our data on adaptation to position-dependent viscous forces (Figure 2) and previous data on generalization across large displacements (Shadmehr and Moussavi 2000), we considered both Gaussian and linear encoding of position space. We first assumed a Gaussian coding of limb position space and assessed the optimal width of basis elements to fit the data. We were surprised to find that the optimal full width at half-maximum of each element was approximately 80 cm (standard deviation of Gaussian function, $\sigma = 34$ cm). This very broad tuning of position space by Gaussian basis elements formed an essentially linear position-dependent receptive field over a workspace three times the width of the training space. Because this model produced essentially monotonic encoding of position throughout our training space and beyond, we decided to study in detail a model with simple linear position encoding. Indeed, studies of the motor cortex (Georgopoulos et al. 1984; Sergio and Kalaska 1997), somatosensory cortex (Prud'homme and Kalaska 1994; Tillery et al. 1996), and spinocerebellar tract (Bosco et al. 1996) have found that cells in these areas code limb static-position globally and often linearly.

One way to represent limb position and velocity is with basis elements that encode each variable and then add them. However, additive encoding cannot adapt to fields that are nonlinear functions of position and velocity, e.g., $f(x, \dot{x}) = (x/d) \cdot B\dot{x}$. This is the force field that describes the task we considered in the previous section. A theoretical study suggests that to adapt to such nonlinear fields, the basis functions of the combined space must be formed multiplicatively rather than additively (Pouget and Sejnowski 1997). We chose to use a multiplicative combination of
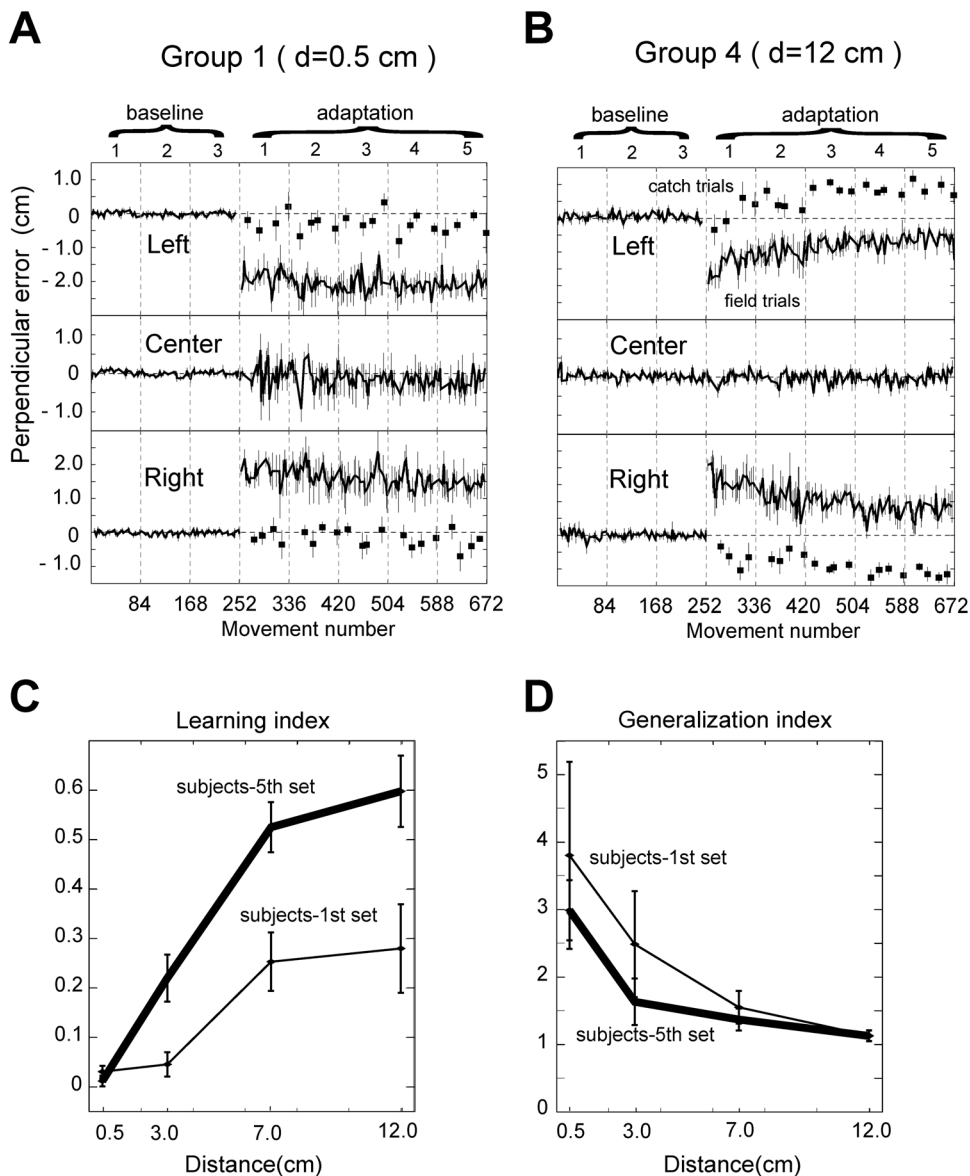
Figure 2. Movement Errors and Learning Performance as a Function of the Separation Distance

(A) PE averaged across six subjects of Group 1 ($d = 0.5$ cm). Squares indicate catch trials. Error bars show SEM. The average SEM at the center is 1.3 mm in the baseline sets and 6.6 mm in the adaptation sets for Group 1. The average SEM is 2.1 mm in the baseline sets and 2.4 mm in the adaptation sets for Group 4.
(B) Errors were averaged across six subjects of Group 4 ($d = 12$ cm).
(C) Average learning index (Equation 1) across groups. Learning index is plotted against the separation distance between movements. Thin lines show the first adaptation set and thick lines show the last adaptation set. $n = 6$ for each distance. Error bars show SEM.
(D) Generalization index (Equation 2) against spatial distance between movements.
DOI: 10.1371/journal.pbio.0000025.g002

position and velocity. Thus, we hypothesized that position and velocity encoding are combined via a gain-field mechanism; i.e., the bases have velocity-dependent receptive fields, and the discharges in these receptive fields are linearly modulated by arm position (Figure 3A).

We found that when a network learned to represent the force field via a gain-field encoding of limb position and velocity, it produced movements that matched the generalization pattern both in the current experiment and in earlier reports. Figure 3C and 3D shows the time series of errors from the simulated controller in the same format as in Figure

2. The learning and generalization indexes are plotted in Figure 3E and 3F, and when compared to the values calculated from subject data, they show good agreement (correlation coefficient = 0.96 for the learning index and 0.99 for generalization).

Why Does Gain-Field Coding Account for the Patterns of Generalization?

We used the data in Figure 2D (which shows generalization as a function of spatial location) to estimate the position sensitivity of the basis elements. To explain how this works, we illustrated the process using a model that has only two
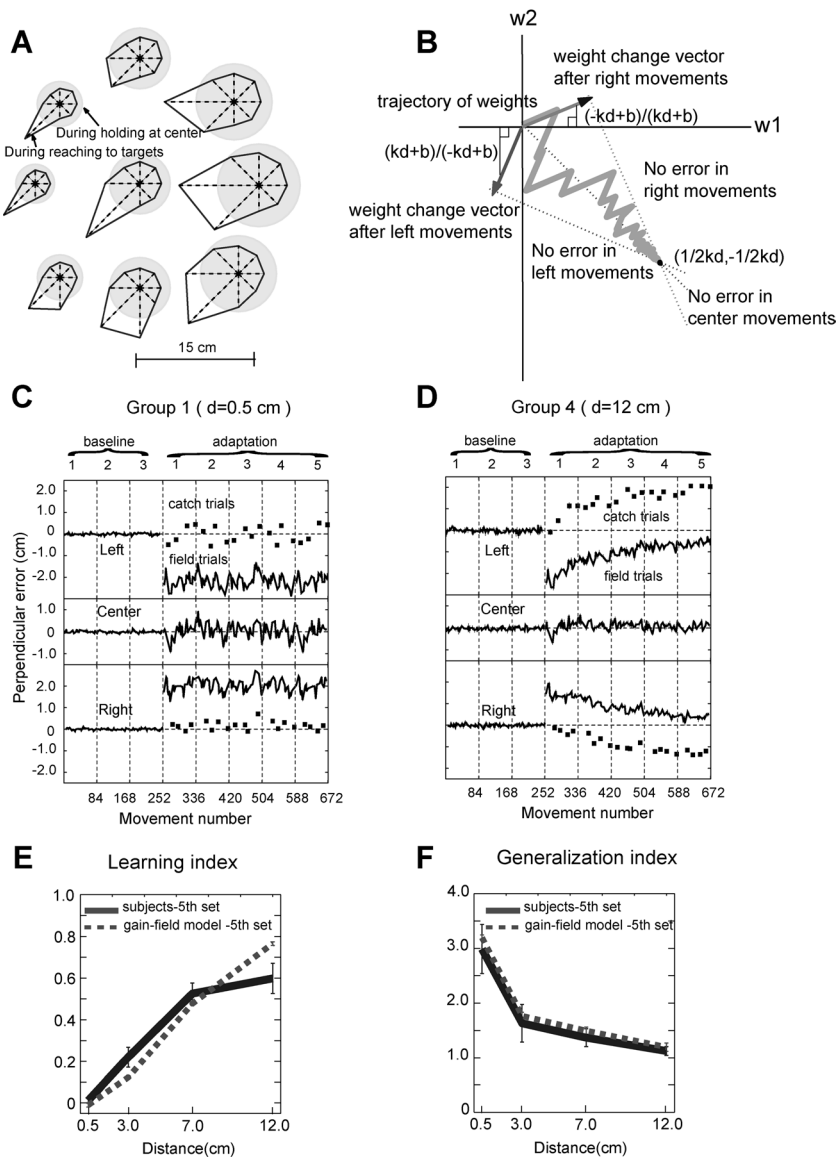
Figure 3. Adaptation with Basis Elements That Encode Limb Position and Velocity as a Gain-Field

(A) A polar plot of activation pattern for a typical basis function in the model. The polar plot at the center represents activation for an eight-direction center-out reaching task (targets at 10 cm). Starting point of each movement is the center of the polar plot. The shaded circle represents the activation during a center-hold period and the polygon represents average activation during the movement period. The eight polar plots on the periphery represent activation for eight different starting positions. Each starting position corresponds to the location of the center of each polar plot. The preferred positional gradient of this particular basis function has a rightward direction. The preferred velocity is an elbow flexion at 62°/s.

(B) A state diagram of weights in a simple system with two basis functions. The trajectory from the origin to ($\frac{1}{2}kd$, $-\frac{1}{2}kd$) shows how weights converge to the final values trial-by-trial. Three dotted lines represent weights for no errors on the left, center, and right movements, respectively. Two vectors represent the direction of weight change after left and right movements each.

(C) The bases were used in an adaptive controller to learn the task in Figure 1. Format is the same as Figure 2A; correlation coefficient of the simulated to subject data is 0.97.

(D) Simulated movement errors in an experiment where spatial distance was the same as in Group 4 in Figure 2B; correlation coefficient is 0.86.

(E) Learning index of the last target set against spatial distance. Dotted lines are from the simulation and thick solid lines are from subjects; correlation coefficient is 0.96. Note that thick solid lines are the same lines as in Figure 2C.

(F) Generalization index in the last target set against spatial distance. Dotted lines are from the simulation and solid lines are from subjects; correlation coefficient is 0.99.

basis elements and where the basis elements only encode position. We can limit ourselves to two basis elements because the position dependence of our task was restricted to a single dimension. Similarly, because our task only required movements in one direction, we can explain the behavior of the

model without including velocity sensitivity. However, while the reduced model is useful in describing the basic principles, we fit the full model to the data. This was for two reasons. First, the actual adaptation requires velocity coding. Second, the full model is an extension of models used previously

(Thoroughman and Shadmehr 2000; Donchin and Shadmehr 2002), and we wanted to be sure that the new model could account for all available datasets.

We chose to use hand position (x) for a simpler description of the model in this section. However, in the full model, the coordinate system of limb position is in terms of joint angles. In the simple model, the two bases are $g_1(x) = kx + b$ and $g_2(x) = -kx + b$, where x is hand position, k is the sensitivity of the element's output to changes in hand position, and b is the constant. The net expected force is a weighted sum of these two functions, $\hat{f}(x) = w_1 \cdot g_1(x) + w_2 \cdot g_2(x)$, , where w_1 and w_2 are weights for g_1 and g_2, respectively (refer to Materials and Methods for details). After each trial, weights were updated so that the expected force function ultimately approximated the externally applied force. The applied forces in the previous section were a linear function of hand position, e.g., $f(x) = x/d$, where d is the separation distance. When $w_1 = \frac{1}{2}dk$ and $w_2 = -\frac{1}{2}dk$, this field is perfectly approximated by the bases. Thus, with training, weights were updated to converge to these values. The weight trajectory from the initial value to the final value is drawn on the state-space diagram in Figure 3B. Three dotted lines represent weights where the expected force is correct for right, middle, and left movements, respectively, and the intersecting point of these three lines is the correct weight for all three positions, i.e., the final value. The amount of weight change after each trial depended on the force error experienced on that trial and the activation of each basis element on that trial. Thus, after movements on the right $(x = d)$, w_1 changed more than w_2 because g_1 was bigger than g_2 for movements on the right, and after movements on the left $(x = -d)$, the opposite was true. The diagram shows the weight change after right and left movements, respectively, where the slope of each vector $(\Delta w_2 / \Delta w_1)$ is equal to the ratio $g_2(x)/g_1(x)$; e.g., the slope of the vector is $(-kd + b)/(kd + b)$ after right movements and $(kd + b)/(-kd + b)$ after left movements. Fast learning occurred if these vectors were closely aligned to the middle dotted line (slope $= -1$) because weights can follow a shorter path from the initial to final values. Thus, we can see from the simple model that a larger k, a larger d, or a smaller b will produce faster learning.

The trial-by-trial variation in the center movements is also clarified by examining this state-space diagram; i.e., any deviation from the middle dotted line means a nonzero force expectation for the center movements and larger deviations correspond to larger errors in the center movements. Thus, update vectors with a slope near -1 lead to both faster learning and smaller variance in the middle movements. Therefore, the slope k and the constant b are important parameters in our model that determine the learning rate and the generalization to the center movements. For a given k and b, the separation distance d will modulate the learning rate and generalization.

We adjusted k and b for the bases to fit the simulation's performance to the generalization observed in subjects (see Figure 2D). Figure 3C–3F shows good agreement between our simulation and subject data. However, the question is, can this same model with the same parameter values explain other behavioral data?
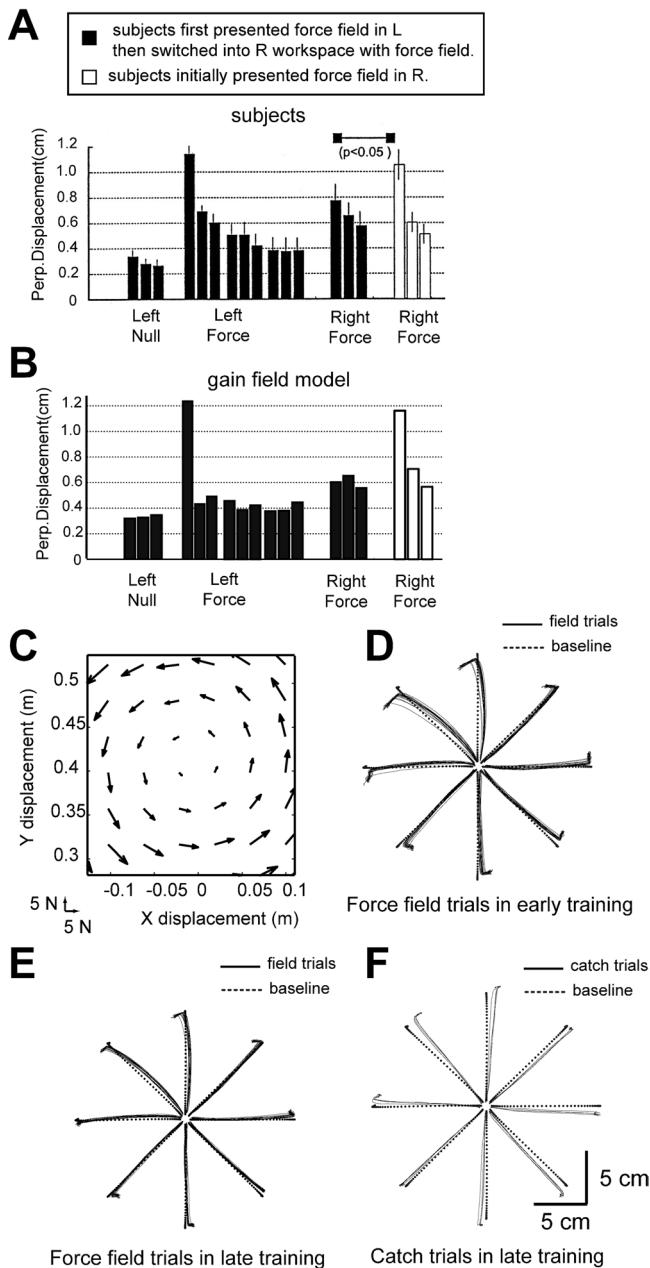
Testing the Model on Previously Published Results

We found that the gain-field model could also account for a number of other previously published results. The experi-ments we focus on here are adapting to a field that depended only on limb velocity and not position (Shadmehr and Moussavi 2000) and adapting to a field that depended only on limb position and not velocity (Flash and Gurevich 1997). We found that exactly the same basis elements that fit the generalization pattern in Figure 2D also accounted for behavior in these paradigms.

Shadmehr and Moussavi (2000) present an experiment where subjects trained in a position-independent field in one workspace (hand to the far left) and were then tested in a different workspace (hand to the far right). Essentially, the question was, if field $f = 1$ was presented at position $x = d$, what kind of force would be expected at another location? We can predict what will happen by analyzing the reduced model. All the weights that satisfy $w_1(kd + b) + w_2(-kd + b) = 1$ can approximate this force field correctly. However, the slope of the weight change vector, i.e., $(-kd + b)/(kd + b)$, will determine the final weights uniquely as $w_1 = (kd + b)/(2k^2d^2 + 2b^2)$ and $w_2 = (-kd + b)/(2k^2d^2 + 2b^2)$. Thus, the force function approximated by these weights is $\hat{f}(x) = (k^2d \cdot x)/(k^2d^2 + b^2) + b^2/(k^2d^2 + b^2)$. Therefore, the expected force is again a linear function with slope $(k^2d)/(k^2d^2 + b^2)$. This slope decreases as the gain k decreases and the constant b increases. Importantly, when the trained position d is close to the zero of the coordinate axis, the slope is also close to zero, making the generalization function flat (i.e., global generalization).

In the Shadmehr and Moussavi experiment (2000), subjects trained in a clockwise viscous curl-field ($F = B \cdot \dot{x}$, $B = [0 - 13; 13\ 0]$ N m/s) in the "left workspace" (shoulder in a flexed posture) and were then tested in the same field with the hand in a workspace 80 cm to the right (shoulder in an extended posture). The idea was to see whether there is any general-ization of learning from the left workspace to the right workspace. Since the viscous curl-field perturbed the subject's hand in the perpendicular direction of hand velocity, as a movement error, they measured the maximum perpendicular displacement from the straight line connecting start and target position. For a direct comparison, we used the same measure, maximum PE, shown in Figure 4. During the left workspace training, the movement errors decreased (Figure 4A). After the training on the left workspace, Shadmehr and Moussavi (2000) tested these subjects in the right workspace. Their errors on the right are significantly smaller than the errors of control subjects, who did not train on the left, indicating a transfer of adaptation from the left workspace to right (Figure 4A). When confronted with the same protocol, the bases that had fit out data in Figure 3 produced a pattern of generalization across large distances that was quite similar to that of the subjects' generalization (Figure 4B; correlation coefficient $= 0.89$).

It is also known that subjects can adapt to position-dependent spring-like force fields (Flash and Gurevich 1997; Tong et al. 2002), where force increases linearly with hand displacement. Flash and Gurevich (1997) showed that immediately after the introduction of the force field, the movement trajectories deviated from straight hand paths with large endpoint errors. However, with training, move-ments became straighter and errors decreased. This was not simply a result of increased stiffness because, when the force field was removed unexpectedly, the trajectories and end-point errors were on the opposite side of those in the early force field, indicating a proper internal model. Figure 4C

A

B

C

D

Force field trials in early training

E

Force field trials in late training

F

Catch trials in late training

Figure 4. Gain-Field Representation Reproduces Previously Reported Patterns of Spatial Generalization

(A) Figure 4A from Shadmehr and Moussavi (2000). The right workspace is separated from the left workspace by 80 cm. Each histogram bar is average maximum PE of 64 movements. Smaller errors of trained group than those of control group at right workspace indicate the transfer of learning from left to right workspace.
(B) Simulation results in the same format as (A); correlation coefficient to subject data is 0.89.
(C) A spring-like force field, $\vec{F} = K \cdot \vec{x}$ ($K = [0\ -55; 55\ 0]$ N/m), was used for simulation.
(D) Hand path trajectories in field trials from the first set in the spring-like force field. Each set consists of 192 center-out movements. Targets are given at eight positions on 10 cm circumference in pseudorandom order.
(E) Hand path trajectories in field trials from the fifth set of training.
(F) Hand path trajectories of catch trials from the fifth set.
DOI: 10.1371/journal.pbio.0000025.g004

shows a position-dependent field in this category, and Figure 4D–4F shows movements made by the simulation as it adapts to this field. This pattern of adaptation is similar to reported values in human data (Flash and Gurevich 1997).

## Testing the Model's Predictions

Our hypothesis regarding adaptation with a basis that encodes position and velocity as a linear gain-field has two interesting consequences: (1) a change to the pattern of forces can substantially increase the difficulty of a task; and (2) there should be hypergeneralization; i.e., forces expected in an untrained part of the workspace may be larger than the ones experienced at a trained location.

Figure 5A shows the pattern of forces that was previously shown to be easily adaptable. Figure 5B shows a similar task, where the leftward and rightward forces are separated by the same distance, but instead of making null movements between the left and right positions, null movements are made off to the right of the field trials. The field in Figure 5A is learnable by gain-field basis elements. However, if the internal model is indeed computed with such elements, then for the field in Figure 5B we can make two predictions: (1) this pattern of forces should not be learnable because no linear function can adequately describe this nonlinear pattern of force; and (2) null movements at the "right" should show aftereffects of the center movement despite the fact that no forces are ever present.

We tested these predictions in two separate groups of subjects. Six subjects trained in the force pattern of Figure 5A; their performance is shown in Figure 5C (part of the same data shown in Figure 2C and 2D). Five subjects trained in the force pattern of Figure 5B; their performance is shown in Figure 5C. As the model predicted, the performance of subjects in the forces of Figure 5B was significantly worse than in forces of Figure 5A (paired t-test on average learning index across sets: $t = 2.51$, d.f. $= 9$, $p < 0.05$). Recall that movements at the "right" were always in the null field. However, as the model predicted, there was significant generalization here since these movements were significantly biased to the left (t-test: $t = -8.13$, d.f. $= 4$, $p < 0.001$).

An interesting property of systems that learn with gain-fields is that in some conditions, local adaptation should result in an increasing extrapolation, i.e., hypergeneralization. Consider a situation in which, during the training sets, subjects make movements in the center as well as at 5 cm to the right (Figure 5D). A counterclockwise curl-field is applied to the center movements, while no forces are applied to movements at right. Because the coding of limb position is linear in the gain-fields, the approximated force function is a linear function that grows from right to left; i.e., the adaptive system should expect larger forces when movements are to the left of center. We tested this in four subjects. During the test set, catch trials were introduced for center movements and occasionally a movement was performed at left (Figure 5D). These movements at the left were always in a null field. The catch trials of left movements were significantly larger than those of center movements, which is consistent with the prediction of hypergeneralization (Figure 5E; paired t-test: $t = 4.35$, d.f. $= 6$, $p < 0.005$).

One concern is the weak learning during the training sets. However, the average learning index for the center movements in the last set is 0.46, and this is significantly different
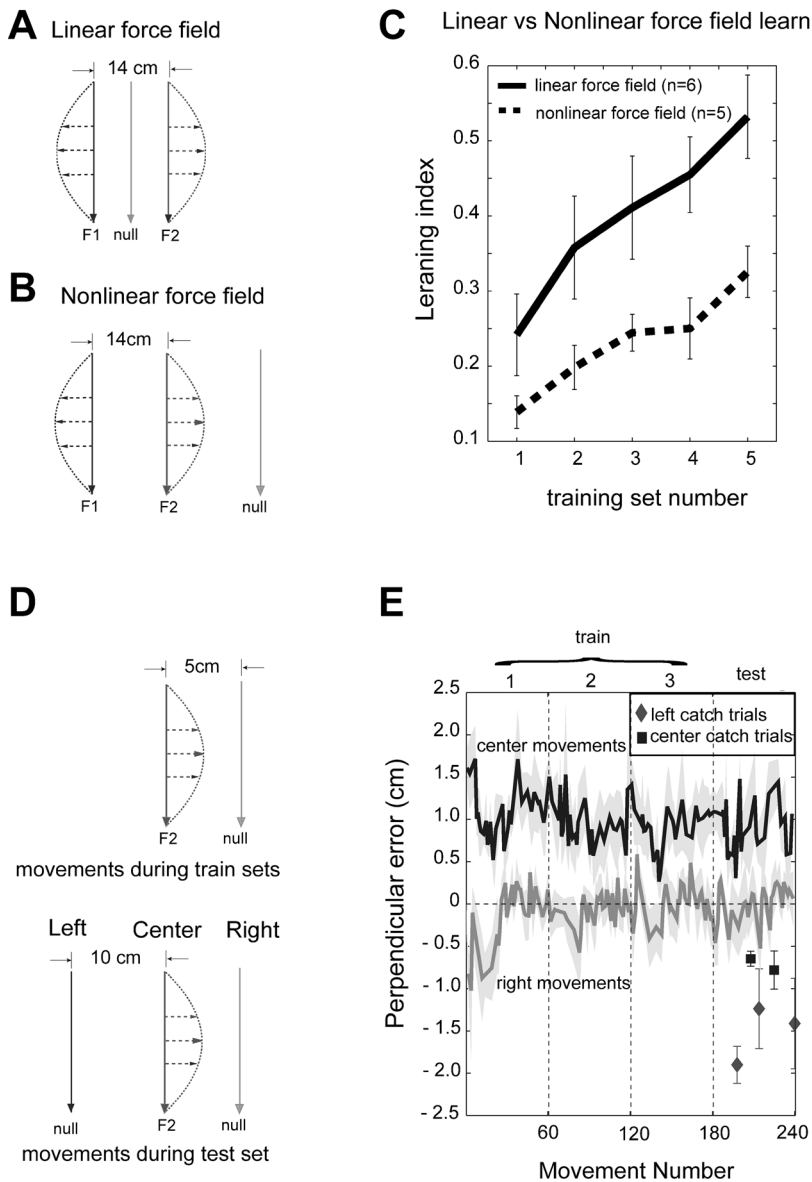
Figure 5. Predictions of the Gain-Field Encoding and Experimental Verification

(A) A field where forces are linearly dependent on both limb position and velocity.
(B) A field where forces are linearly dependent on limb velocity but nonlinearly dependent on limb position. Gain-field encoding predicts that the field in (B) will be harder to learn than one in (A).
(C) Learning index of subjects ($n = 6$) for the paradigm in (A) and subjects ($n = 5$) for the paradigm in (B).
(D) Gain-field encoding predicts hypergeneralization. The figure shows movements and its associated force field during training and test sets.
(E) Performance of subjects ($n = 4$) for the paradigm in (D). Dark lines are errors in center movements and gray lines are errors in right movements. The shaded areas represent the SEM. Filled diamonds show the catch trials for the left movements during test set; filled squares show the catch trials for center movements.
DOI: 10.1371/journal.pbio.0000025.g005

from zero (*t*-test: $t = 3.32$, d.f. $= 3$, $p < 0.05$). Considering that the separation distance between null and force field movements is only 5 cm, this learning index is consistent with the learning index curve in Figure 2C and consistent with the learning possible with the proposed bases. More importantly, despite this small learning in the training space, we found significantly bigger aftereffects in the test space, i.e., hypergeneralization. Our simulations suggest that this hypergeneralization could not be due to varying limb inertia and/or stiffness as a function of limb position.

Discussion

When people reach to various directions in a small workspace, velocity- or acceleration-dependent forces that they experience are generalized broadly to other arm positions as far as 80 cm away (Shadmehr and Moussavi 2000). These results argue that the neural elements with which the brain represents the dynamics of reaching movements may not be very sensitive to limb position, in contrast to findings that they are quite sensitive to velocity and acceleration. Lack of sensitivity would explain extensive

generalization. However, it is known that humans can adapt reaching movements to position-dependent spring-like forces (Flash and Gurevich 1997). The ability to adapt to position-dependent fields suggests that the internal model can have steep position sensitivity (Tong et al. 2002). This apparent contradiction raised doubts about our understanding of representation and generalization of limb position. Therefore, we closely examined patterns of generalization as a function of limb position and asked whether these results could be explained by a single representation.

We hypothesized that adaptation to arm dynamics was due to an internal model that computed a map from sensory variables (limb position and velocity) to motor commands (force or torque). These elements were sensitive to both the position and velocity of the arm. The main question was how these variables were encoded. We first performed behavioral experiments to characterize the limits of adaptation to position-dependent forces. This allowed us to quantify the sensitivity of position coding. We found that generalization to neighboring movements decayed gradually with separation distance, implying a very broad position encoding. We found that a Gaussian representation would require a full width at half-maximum of approximately 80 cm to explain our results. Since a Gaussian this broad would be indistinguishable from a monotonic function, we used a linear function instead. A linear basis is a simple monotonic encoding of position space. We combined position and velocity by making position a linear gain-field on the directional sensitivity.

Using a gain-field basis to simulate the learning of arm movements, we found that the parameters that fit our pattern of decaying generalization could also account for a number of previously published results on generalization in position space. These results were the generalization of learning over a large workspace and the ability to learn stiffness fields. Additionally, we tested two behavioral predictions of our model to further test the hypothesis of gain-field coding. Theory predicted that a simple rearrangement of position-dependent forces would change a task from easily learnable to very difficult. It also predicted that in a two-point adaptation paradigm, expected forces would be extrapolated so that larger forces would be expected outside the trained workspace. The behavioral results agreed with these theoretical predictions. Thus, our model used a multiplicative interaction between coding of limb position and velocity to explain behavioral data during learning dynamics of reaching movements and successfully predicted data from a variety of experimental paradigms.

Learning by Population Coding via Gain-Fields vs. Modular Decomposition

Ghahramani and Wolpert (1997) studied similar starting-position-dependent visuomotor mappings in which two opposite visual perturbations were applied to the two starting positions of movements. Subjects learned the two starting-point-dependent visuomotor mappings and generalized this learning to intermediate starting positions using interpolation. Their interpretation of this result was that the brain employs two visuomotor experts, each of which is responsible for one of the two visuomotor mappings, and interpolates to intermediate starting locations using a weighted average of the two experts. If the brain employs this modular decomposition strategy in learning dynamics as well, two of our

findings will be hard to explain: (1) learning ability changes with the separation distance between starting positions, although only three experts are required regardless of the separation distance; (2) learning ability decreases when the force field pattern is nonlinear, although the same number of experts (three) is required for both linear and nonlinear patterns.

However, if the internal models for dynamics are represented as a population code with gain-fields, these two factors are easily explained by the proximity of the population code for the close distance and the monotonic change of population code with the starting positions. Gribble and Scott (2002) examined cell responses for three different dynamics conditions: the elbow-joint-dependent viscous curl-field, the shoulder-joint-dependent viscous curl-field, and both joint-dependent viscous curl-fields. In all three conditions, monkeys were trained to the level that the kinematic properties of movements were close to the baseline. Gribble and Scott (2002) found that many cells that responded to one joint-dependent field also responded continuously to the other joint-dependent field, supporting a single controller hypothesis with population coding rather than separate experts. Therefore, the available data on learning of internal models of dynamics seem to be inconsistent with modular decomposition.

Gain-Field Coding of Position and Velocity

Multiplicative interaction of two independent variables in cell encoding is called gain-field coding. Although we described our gain-field as a velocity-dependent signal that is modulated by limb position, it can also be described as a position-dependent signal that gets modulated by limb velocity. Gain-fields originally described the tuning properties of cells that are responsive to both visual stimuli and eye position in area 7a of the parietal cortex. The receptive field of these cells remains retinotopic while the gain of the retinotopic curve is modulated linearly by eye position (Andersen et al. 1985). This multiplicative response is computationally advantageous because a population of such cells provides a complete basis set for the combined space (Poggio and Girosi 1990; Pouget and Sejnowski 1997); i.e., any arbitrary function, linear or nonlinear, in the combined space can be approximated as the weighted sum of these basis elements. Considering that many computational problems in the motor system use both direction of reaching movement and hand position information, it seems attractive to have a complete set of basis functions encoding these two variables. The behavior that we recorded from our subjects is in agreement with the patterns of interference that such bases would produce.

There are other ways to form a basis set. A prominent example is an additive basis set. In an additive set, a function of position is added to a function of movement direction. Some neurophysiological experiments have used this kind of model, rather than a multiplicative model, to relate neural discharge in the motor cortex and cerebellum to limb position and velocity (Fu et al. 1997a, 1997b). However, if limb position and velocity are coded additively, a population of such basis elements cannot approximate the force fields that our subjects learned, as shown, for example, in Figure 2B. With an additive basis set, one cannot approximate functions that include a nonlinear interaction between two indepen-

dent variables, such as multiplication, even if each basis set before combining is complete for each independent subspace (Pouget and Sejnowski 1997). Simulation results using additive basis elements corroborate this argument, as these simulations showed much less learning than our subjects (data not shown).

Neurophysiological Findings Related to Our Model

Although our computational model was derived from psychophysical experiments, a number of neurophysiological findings seem to be consistent with properties of our basis elements.

First, neurophysiological recordings support our monotonic position encoding in joint angle coordinates. Human muscle spindle afferents, both individually and as a population, represent static joint position monotonically (Cordo et al. 2002). This monotonic position encoding, possibly originating from the property of such peripheral afferents, is consistently found in the central nervous system as well. Georgopoulos et al. (1984) found that the steady-state discharge rate of cells in the motor cortex and area 5 varied with the static position of the hand in two-dimensional space, and the neuronal response surface was described by a plane, indicating that individual cells in these areas encode position of the hand monotonically (and continuously) in space (Prud'homme and Kalaska 1994; Sergio and Kalaska 1997). This monotonic response to hand or foot position was also observed in S1 of primates and spinocerebellar neurons of cats (Bosco et al. 1996; Tillery et al. 1996). However, in those studies, it is unclear which variable between hand position and joint angle is an independent input to these cell responses since these two variables are almost linearly related in a small workspace. Scott et al. elucidated this point, showing that neural activity in parietal and motor cortical cells changed when the hand was maintained at the same location but with two different arm orientations (Scott and Kalaska 1997; Scott et al. 1997); i.e., at least parts of the motor cortex seem to encode limb position in joint angle coordinates rather than hand-based Cartesian coordinates.

Another distinct property of our basis elements is that their activity is modulated by both position and velocity. Caminiti et al. (1990) found that as movements with similar trajectories were made within different parts of space, some motor cortical cells' preferred directions changed spatial orientation, indicating that they encoded direction of movement in a way that was dependent on the position of the arm in space. Similar interaction between movement direction and arm posture, wherein cells were directionally tuned but the overall activity levels varied with arm postures, was found in S1 (Prud'homme and Kalaska 1994). Sergio and Kalaska (1997) studied this interaction more systematically during the static isometric force generation. They found that the overall level of tonic activity of M1 cells varied monotonically with the hand position, and the preferred direction tended to rotate with the hand position in an arc-like pattern. All of this is reflected in the gain-field representation of the bases shown in Figure 3A.

Lastly, the output of our basis elements is associated with a preferred joint torque vector. With adaptation, these torques rotate. Prud'homme and Kalaska (1994) found that the discharge of M1 cells changed when the monkey compensated for inertial loads. Li et al. (2001) also found similar load-dependent activity changes during adaptation, and for the entire neuronal population, the shift in preferred direction of M1 cells matched the shift observed for muscles. Similar studies by Gribble and Scott (2002) support the idea that the output of elements representing internal models is related to joint torques.

Monotonic Position Encoding

Although we used a linear encoding of limb position because of its mathematical simplicity, data based on three positions are not sufficient to distinguish a linear from a nonlinear basis function. Therefore, at this point, our basis functions are best viewed as having a monotonic property. Monotonic gain-field coding of position and velocity makes an intriguing prediction regarding behavior. When two different forces are experienced at two different arm positions (as in Figure 5A), the generalization function is a linear function that connects the two forces at the two positions. Thus, the forces could grow outside the trained workspace. We observed this strange feature of linear gain-fields in the experiment shown in Figure 5D. In that experiment, we found that the aftereffects in the movements at a new workspace (left side) were larger than the aftereffects in the trained space (center). This suggests that larger forces were expected at left, an example of hypergeneralization. This raises the concern that with gain-fields, generalization in two-point adaptation might keep growing. However, considering that the reach workspace is bounded and the gain change by the position is very gradual, significantly larger generalization occurs only when trained force fields are specifically position dependent (Figure 5D).

Explicit Cue Learning

Another issue is that our findings may be the result of limits of visual acuity; i.e., a decreased ability to distinguish the starting positions might cause a spurious finding of position-dependent coding. One way we addressed this concern was to use color-coding to make sure subjects could distinguish the left, center, and right targets. However, it is possible that the system is not capable of using color cues while it is capable of using spatial cues. Note that this interpretation implies that, at large separations, position serves as an explicit cue triggering separate internal models. That interpretation is not consistent with the earlier results in which generalization across large distances seemed to imply that position is represented continuously. If position is a discrete cue for building different internal models, then it is not clear how one could learn a force field that depends continuously on position, as in spring-like fields. It is also not clear why learning ability would decrease in the nonlinear force field pattern. Thus, although it is possible that some effects are due to the explicit cues, this cannot entirely explain our findings without a continuous encoding of position.

In sum, we report that generalization properties of learning arm dynamics can be explained using basis elements that encode limb position and velocity in intrinsic coordinates using a multiplicative, gain-field interaction. Hand position seems to be encoded monotonically and velocity seems to be encoded using Gaussian elements. The result is a gain-field where position monotonically modulates the gain of velocity tuning. We predict that this encoding will be

reflected in the activity of neurons responsible for adaptation to dynamics of reaching movements.

Materials and Methods

Subjects. Thirty-three healthy individuals (16 women and 17 men) participated in this study. The average age was 27.5 y (range: 21–50 y). The study protocol was approved by the Johns Hopkins University School of Medicine Institutional Review Board and all subjects signed a consent form.

Performance measures. As a measure of error, we report the displacement perpendicular to target direction at 250 ms into the movement (PE). However, we also tried other measures, such as perpendicular displacement at the maximum tangential velocity, maximum perpendicular displacement, and averaged perpendicular displacement during early phase of movement. Results that we present here are consistent among all these measures of error, and we report only PE.

During adaptation, trajectories in field trials become straighter, while the trajectories of catch trials become approximately a mirror image of those in earlier field movements (Shadmehr and Mussa-Ivaldi 1994). Therefore, the PE of field trials decreases and the PE of catch trials increases. Based on this observation, we quantified a learning index:

$$Learning\ Index = \frac{p.e._{catch\ trial}}{p.e._{catch\ trial} - p.e._{field\ trial}} \quad (1)$$

We quantified the effect of error experienced in one movement on another movement as a function of their spatial distance. In the experiment outlined in Figure 1, there were never any forces during the center movement. The error experienced in a neighboring movement would cause a change in the subsequent movement at center. This change results in increased variance of errors at center. Therefore, a measure of generalization of error is the ratio of variance of error in trials where forces were present to the left and right of the center movement, to variance of error in baseline trials when no forces were present:

$$Generalization\ Index = \frac{standard\ deviation\ of\ p.e._{adaptation}}{standard\ deviation\ of\ p.e._{baseline}} \quad (2)$$

Computational modeling. The internal model may be computed as a population coding via a set of basis elements, each encoding some aspect of the limb's state (Donchin and Shadmehr 2002). Neurophysiological studies show that in tasks similar to the current paradigm, the preferred direction of cells tend to change during adaptation (Li et al. 2001; Gribble and Scott 2002). In our model, we assumed a "preferred" torque vector is associated with each basis. With training, the preferred torque vectors change, resulting in a more accurate representation of the force field. The internal model is:

$$\hat{\tau}_{env} = \sum_i w_i \cdot g_i(\underline{z}) \quad (3)$$

where $\hat{\tau}_{env}$ is the expected environmental torque, $\underline{z}$ is a desired state of the limb (consisting of limb position and velocity), g_i is a basis element, and w_i is a torque vector composed of shoulder torque and elbow torque, corresponding to each basis element. In training, adaptation is realized by a trial-to-trial update of torque vectors following gradient descent in order to decrease the difference between the actual torque experienced during the movement and the expectation of torque currently predicted by the internal model.

We hypothesized that the bases have a receptive field in terms of the arm's velocity (in joint space) and that the discharge at this receptive field is modulated monotonically as a function of the arm's position; i.e., the elements represent the arm's position and velocity as a gain-field:

$$g\left(\underline{q}, \underline{\dot{q}}\right) = g_{position,i}\left(\underline{q}\right) \cdot g_{velocity,i}\left(\underline{\dot{q}}\right)$$

$$g_{positions,i}\left(\underline{q}\right) = k^T \cdot \underline{q} + b$$

$$g_{velocity,i}\left(\underline{\dot{q}}\right) = \exp\left(\|\underline{\dot{q}} - \underline{\dot{q}_i}\|^2 / 2\sigma^2\right) \quad (4)$$

Typical output of this basis for various limb positions and movement directions is plotted in Figure 3A. The position-dependent term is a linear function that encodes joint angles, $\underline{q} = (\theta_{shoulder}, \theta_{elbow})$, while the velocity-dependent term encodes joint velocities. The choice of intrinsic rather than extrinsic coordinates is important because the extrinsic representation of limb velocity cannot account for behavioral data on patterns of generalization (Shadmehr and Moussavi 2000; Malfait et al. 2002). The gradient vector k reflects sensitivities for the shoulder and elbow displacement, and b is a constant. The direction of gradient vectors is uniformly distributed in joint angle space with 45° increments. A basis function with 0° direction of gradient is sensitive only to the shoulder angle changes, whereas a basis with a gradient in 90° direction is sensitive only to elbow angle changes. The velocity-dependent term is a Gaussian function encoding joint angular velocity ($\dot{q}_d$, a 2 × 1 vector composed of shoulder and elbow joint velocity) centered on the preferred velocity ($\dot{q}_i$).

To fit experimental data, we varied only two parameters of the model: the slope (k, magnitude of a gradient vector) and the constant (b) in Equation 4. All the other parameters were fixed in the following manner: (1) the directions of gradient vectors were uniformly distributed from 0° to 315° with a 45° increment; (2) the preferred velocities are uniformly tiled in joint velocity space with a 20.6°/s spacing and width; (3) the total number of basis elements was equal to the number of the preferred positional gradients multiplied by the number of preferred velocities because we used every possible combination of gradient and preferred velocity; (4) random noise was injected into the torque in the simulated system so that movements in the null field had the same standard deviation of PE as did the subjects' movements. In exploring the parameter space of k and b, we found that 1 rad^{-1} and 1.3 for the slope and constant, respectively, gave a good fit of generalization as a function of separation distance. To simulate human arm reaching, we used a model of the arm's dynamics that described the physics of our experimental setup (Shadmehr and Mussa-Ivaldi 1994).

Supporting Information

Dataset S1. Trial-by-Trial Analysis
View online at DOI: 10.1371/journal.pbio.0000025.sd001 (60 KB PDF).

References

Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. Science 230: 456–458.

Bosco G, Rankin A, Poppele RE (1996) Representation of passive hindlimb postures in cat spinocerebellar activity. J Neurophysiol 76: 715–726.

Caminiti R, Johnson PB, Urbano A (1990) Making arm movements within different parts of space: Dynamic aspects in the primate motor cortex. J Neurosci 10: 2039–2058.

Conditt MA, Mussa-Ivaldi FA (1999) Central representation of time during motor learning. Proc Natl Acad Sci U S A 96: 11625–11630.

Conditt MA, Gandolfo F, Mussa-Ivaldi FA (1997) The motor system does not learn the dynamics of the arm by rote memorization of past experience. J Neurophysiol 78: 554–560.

Cordo P, Flores-Vieira C, Verschueren MPS, Inglis JT, Gurfinkel V (2002) Position sensitivity of human muscle spindles: Single afferent and population representation. J Neurophysiol 87: 1186–1195.

Donchin O, Shadmehr R. (2002) Linking motor learning to function approximation: Learning in an unlearnable force field. In: Diettrich TG, Becker S, Ghahramani Z, editors. Advances in Neural Information Processing Systems 14.Cambridge, MA: MIT Press. pp. 197–203.

Flash T, Gurevich I (1997) Models of motor adaptation and impedance control in human arm movements. In: Morasso P, Sanguineti V, editors. Self-organization, computational maps, and motor control. Amsterdam: Elsevier Science. pp. 423–481.

Fu QG, Flament D, Coltz JD, Ebner TJ (1997a) Relationship of cerebellar Purkinje cell simple spike discharge to movement kinematics in the monkey. J Neurophysiol 78: 478–491.

Fu QG, Mason CR, Flament D, Coltz JD, Ebner TJ (1997b) Movement kinematics encoded in complex spike discharge of primate cerebellar Purkinje cells. Neuroreport 8: 523–529.

Georgopoulos AP, Caminiti R, Kalaska JF (1984) Static spatial effects in motor cortex and area 5: Quantitative relations in a two-dimensional space. Exp Brain Res 54: 446–454.

Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neural population coding of movement direction. Science 233: 1416–1419.

Ghahramani Z, Wolpert DM (1997) Modular decomposition in visuomotor learning. Nature 386: 392–395.

Ghez C, Krakauer JW, Sainburg RL, Ghilardi MF (2000) Spatial representation and internal models of limb dynamics in motor learning. In: Gazzaniga MS, editor. The new cognitive neurosciences. Cambridge, MA: MIT Press. pp. 501–514.

Goodbody SJ, Wolpert DM (1998) Temporal and amplitude generalization in motor learning. J Neurophysiol 79: 1825–1838.

Gribble PL, Scott SH (2002) Overlap of internal models in motor cortex for mechanical loads during reaching. Nature 417: 938–941.

Levi R, Camhi JM (2000) Population vector coding by the giant interneurons of the cockroach. J Neurosci 20: 3822–3829.

Li CSR, Padoa-Schioppa C, Bizzi E (2001) Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. Neuron 30: 593–607.

Malfait N, Shiller DM, Ostry DJ (2002) Transfer of motor learning across arm configurations. J Neurosci 22: 9656–9660.

Moran DW, Schwartz AB (1999) Motor cortical representation of speed and direction during reaching. J Neurophysiol 82: 2676–2692.

Poggio T, Girosi F (1990) Regularization algorithms for learning that are equivalent to multilayer networks. Science 247: 978–982.

Pouget A, Sejnowski TJ (1997) Spatial transformations in the parietal cortex using basis functions. J Cogn Neurosci 9: 222–237.

Pouget A, Dayan P, Zemel R (2000) Information processing with population codes. Nat Rev Neurosci 1: 125–132.

Prud'homme MJ, Kalaska JF (1994) Proprioceptive activity in primate primary somatosensory cortex during active arm reaching movements. J Neurophysiol 72: 2280–2301.

Sainburg RL, Ghez C, Kalakanis D (1999) Intersegmental dynamics are controlled by sequential anticipatory, error correction, and postural mechanisms. J Neurophysiol 81: 1045–1056.

Scott SH, Kalaska JF (1997) Reaching movements with similar hand paths but different arm orientation. I. Activity of individual cells in motor cortex. J Neurophysiol 77: 826–852.

Scott SH, Sergio LE, Kalaska JF (1997) Reaching movements with similar hand paths but different arm orientations. II. Activity of individual cells in dorsal premotor cortex and parietal area 5. J Neurophysiol 78: 2413–2426.

Sergio LE, Kalaska JF (1997) Systematic changes in directional tuning of motor cortex cell activity with hand location in the workspace during generation of static isometric forces in constant spatial directions. J Neurophysiol 78: 1170–1174.

Shadmehr R, Moussavi ZMK (2000) Spatial generalization from learning dynamics of reaching movements. J Neurosci 20: 7807–7815.

Shadmehr R, Mussa-Ivaldi FA (1994) Adaptive representation of dynamics during learning of a motor task. J Neurosci 14: 3208–3224.

Steinberg O, Donchin O, Gribova A, Cardoso de Oliveria S, Bergman H, et al. (2002) Neuronal populations in primary motor cortex encode bimanual arm movements. Eur J Neurosci 15: 1371–1380.

Thoroughman KA, Shadmehr R (2000) Learning of action through adaptive combination of motor primitives. Nature 407: 742–747.

Tillery SI, Soechting JF, Ebner TJ (1996) Somatosensory cortical activity in relation to arm posture: Nonuniform spatial tuning. J Neurophysiol 76: 2423–2438.

Tong C, Wolpert DM, Flanagan JR (2002) Kinematics and dynamics are not represented independently in motor working memory: Evidence from an interference study. J Neurosci 22: 1108–1113.