Temporal Evolution of Spatial Computations for Visuomotor Control

David W. Franklin,1,2* Alexandra Reichenbach,2* Sae Franklin,1 and Jörn Diedrichsen2,3
1Computational and Biological Learning Laboratory, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom, 2Motor Control Group, Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom, and 3Brain and Mind Institute, Western University, London, Ontario N6A 5B7, Canada

Goal-directed reaching movements are guided by visual feedback from both target and hand. The classical view is that the brain extracts information about target and hand positions from a visual scene, calculates a difference vector between them, and uses this estimate to control the movement. Here we show that during fast feedback control, this computation is not immediate, but evolves dynamically over time. Immediately after a change in the visual scene, the motor system generates independent responses to the errors in hand and target location. Only about 200 ms later, the changes in target and hand positions are combined appropriately in the response, slowly converging to the true difference vector. Therefore, our results provide evidence for the temporal evolution of spatial computations in the human visuomotor system, in which the accurate difference vector computation is first estimated by a fast approximation.

Key words: motor control; neural computations; visual processing; visuomotor

Significance Statement

The dominant view regarding the neural control of reaching is that the visuomotor system controls movement based on the difference vector—the difference between the positions of the hand and target. We directly test this theory by measuring the responses to visual perturbations over a large range of possible variations in both target and hand displacements. By modeling the nonlinearity of the feedback response, we were able to reveal the temporal evolution of the underlying computations. The visuomotor system first uses an approximation to the difference vector computation, simply combining the nonlinear responses to cursor and target displacements, only arriving at the correct difference vector calculation 200 ms later.

Introduction

Reaching for a glass of water, playing badminton, or shaking someone’s hand—goal-directed movements are a critical part of everyday life. The sensorimotor control system continuously uses visual information about both target and hand to successfully achieve the movement goal (Georgopoulos et al., 1981; Saunders and Knill, 2003, 2004; Dimitriou et al., 2013). According to the classical view, the brain extracts estimates of target and hand positions from a visual scene, calculates a difference vector between them, and uses this signal to compute the required motor commands (Cisek et al., 1998). Given that online motor corrections respond to changing visual information at time lags in the range of 100–120 ms (Day and Lyon, 2000; Franklin and Wolpert, 2008; Reichenbach et al., 2014), these computations have to be executed very quickly.

A complex network of parietal regions contains distinct neural populations encoding target position (Batista et al., 1999), hand position, or velocity (Ashe and Georgopoulos, 1994; Averbeck et al., 2005), as well as their mixture (Buneo et al., 2002), in both eye-centered and hand-centered coordinates (Buneo and Andersen, 2006). Neurophysiological recordings suggest that the two spatial positions are encoded in a neural population code with neurons possessing spatially receptive fields (Duhamel et al., 1997). From this representation, the calculation of a difference vector requires an intermediate layer of neurons with nonlinear gain field combinations between these two inputs (Pouget and Sejnowski, 1997). This problem is structurally equivalent to the
exclusive-or problem, which cannot be solved in a single-layer neuronal network (Minsky and Papert, 1987). The computation of a motor response is additionally complicated by the necessity to integrate signals in retinal, proprioceptive, and motor reference frames (Buneo et al., 2002). Therefore, the network may require some time to converge on the correct difference vector calculation. Hence, we hypothesize that the brain initially uses an approximate computation to generate fast feedback responses when the task goal seems compromised during an ongoing movement. Specifically, we propose that it initially generates corrective responses to hand and target information independently, without integrating them before the output stage (multichannel model).

We use rapid feedback responses as a window into the temporal evolution of the computations integrating the visual signals (Franklin and Wolpert, 2011a; Pruszynski et al., 2011). While participants executed planar reaching movements to a visual target, we displaced the cursor that symbolized the hand, and the target laterally, and measured the fast corrective feedback responses (Fig. 1 A,B). All possible combinations of target and cursor displacements were tested (Fig. 1C). According to the difference vector model (Fig. 2A), the corrective response should be a function of the difference between the target and the cursor displacements (Cisek et al., 1998). However, visual target and cursor estimates might be updated with different speeds (Brenner and Smeets, 2003), or might be weighted differently because the hand estimate is additionally informed by proprioception and efferent signals (van Beers et al., 1999; Sarlegna et al., 2003). Therefore, we additionally tested a more general version of the difference vector model, the weighted difference vector model (Fig. 2B). In contrast, the multichannel model proposes that these visual signals are only integrated at a later stage (Fig. 2C). Because the corrective response to an isolated target or cursor displacement shows a nonlinear saturation with the size of displacement (Fig. 1D,E), all three models make different predictions for the response to a combined displacement (Fig. 2D). These differences allowed us to track the temporal evolution of the corrective response from an approximate response to the fully integrated solution.

Materials and Methods
Subjects
Twenty neurologically healthy, right-handed (Oldfield, 1971) human subjects (15 females) took part in the experiment (mean age, $24.2 \pm 5.7$ years). All subjects were naive as to the purpose of the study and gave their informed consent before participating. Each subject participated in four identical experimental sessions of $\sim 2$ h each. The institutional ethics committee at the University of Cambridge approved the study.

Experimental apparatus and setup
Subjects performed reaching movements to a target while grasping a robotic manipulandum (Fig. 1A). Subjects were seated and restrained in an adjustable chair in front of a robotic rig. The subjects’ right arm rested on an air sled and they grasped the handle of a vBOT robotic interface (Fig. 1A). The vBOT manipulandum is a custom-built planar robotic interface that can measure the position of the handle and generates forces on the hand (Howard et al., 2009). A six-axis force transducer (ATI Nano 25; ATI Industrial Automation) measured the end-point forces applied by the subject at the handle. The position of the vBOT handle was calculated from joint-position sensors (58SA; Industrial Encoders Direct) on the motor axes. Position and force data were sampled at 1 kHz. Visual feedback was provided using a computer monitor mounted above the vBOT and projected into the plane of the movement via a mirror. This virtual reality system, covering the manipulandum, arm, and hand, prevented direct visual feedback of the arm. The exact onset time of any visual stimulus presented to the subject was determined using the video card refresh signal and confirmed using an optical sensor.

Experimental paradigm
Trials were self-paced: the subject initiated each trial by moving the cursor (yellow circle of 1.0 cm diameter representing the subject’s hand position) into the start circle (gray circle of 1.6 cm diameter, which became white once subjects had moved the cursor into the start) located $\sim 20$ cm in front of their chest. Subjects were instructed to maintain their gaze throughout the entire movement on a thin (two point) fixation cross ($1 \times 1$ cm) located at $\pm 5.0$ lateral and 18.75 cm anterior relative to the start position (Fig. 1B). This position was located exactly at the midpoint between the location of the target and cursor displacements and ensured that both occurred at the same retinal eccentricity. The side of the fixation cross (left or right) was switched every block of trials. The movement initiation cue was a slight change in the color of the fixation cross (white to yellow). Once this color change had occurred, subjects were required to initiate their reach within 1000 ms. Subjects made forward-reaching movements from the start circle to a target (yellow circle of 1.0 cm diameter) located 25.0 cm in front of the start position. The movement was considered to have terminated when subjects had maintained the cursor within 1.0 cm of the target for 600 ms. Subjects were then free to return to the start point to initiate the next trial while feedback was provided about the success of the previous trial.

Ideal trials were defined as trials in which the peak speed was between 42 and 58 cm/s, and subjects did not overshoot the target. On these trials, the subjects received positive feedback (e.g., “great” for peak speeds with 46–54 cm/s or “good” for peak speeds between 42 and 46 or 54 and 58 cm/s), and a counter increased by one point. Other messages were provided visually at the end of each trial to inform the subjects of their performance (either “too fast,” “too slow,” or “overshot target”). Overshooting the target was defined as the position of the cursor exceeding the target in the $y$-direction by 1.8 cm or more. All trials were included in the analysis. This feedback was only provided to encourage subjects to produce consistent movements.

To probe corrective responses to changes in the visual feedback of target or hand position, both the target and the cursor were displaced laterally by one of seven possible distances, $[-3, -2, -1, 0, 1, 2, 3]$ cm, when the cursor reached the middle of the movement (12.5 cm from the start). Because target and cursor displacements were independent, this produced a total of 49 different perturbation conditions (Fig. 1C). The zero value indicates that the cursor and/or target was not displaced. Half of the trials were probe trials, in which the visual perturbations were momentary (lasting 250 ms) and the handle of the manipulandum was constrained along a straight line to the target by a physical channel generated by the vBOT. After the perturbation, the displacement of the cursor/target was reversed. Therefore, to reach the target, a response to the perturbation was not needed. The mechanical channel was implemented as a one-dimensional spring with a stiffness of 4000 N·m$^{-1}$ and damping of $2 \cdot N \cdot m^{-1} \cdot s$ acting lateral to the line from the start to the target (Scheidt et al., 2000; Milner and Franklin, 2005). This allowed us to use the force transducer to measure any lateral forces produced against the channel wall in response to a shift in the target or cursor position. This technique is now commonly used to quantify the gain of the visuomotor feedback responses (Franklin and Wolpert, 2008; Dimitriou et al., 2013; Reichenbach et al., 2013, 2014; Franklin et al., 2014), as the size of the restorative force provides a more reliable measure of the early response than the changes in the lateral velocity of the arm movement.

On the other half of the trials, the perturbations were maintained throughout the rest of the trial, and subjects were required to adjust the movement such that the cursor finished in the target. These trials were included to prevent a decrease of the response over the course of the experiment, as found in previous studies (Franklin and Wolpert, 2008; Franklin et al., 2012; Dimitriou et al., 2013). Indeed, the visual feedback responses were maintained throughout the experiments. The average decrease from the first two blocks in the first session to the final five blocks in the fourth session was only $18.9 \pm 10.7\%$, as opposed to a nearly 75% reduction observed when real corrections were never required (Franklin et al., 2012).
Figure 1. Methods and responses to isolated target or cursor perturbations. A, Subjects made reaching movements while grasping a planar robotic manipulandum. Visual feedback of targets and hand position (cursor) were provided in the plane of the reach via a mirror. B, On random probe trials, the cursor, the target, or both were perturbed laterally by one of seven distances for 250 ms while the physical hand position was constrained by a mechanical channel to the straight line between the start and final target. Subjects fixated visual gaze to the fixation cross (yellow), which signaled the start of movement cue by changing color. The presentation side (right or left) for the fixation cross varied from block to block in the experiment. C, The combination of seven possible target perturbations (green) and seven possible cursor perturbations (red) gave rise to 49 separate conditions where the diagonals had identical difference vectors. D, Left, Mean force responses (solid line) and SEM (shaded region) across all subjects to isolated target displacements plotted with respect to the time of displacement onset. The vertical dotted line indicates the mean (± 1 SD) of the response onset time. The mean force responses were averaged over an early interval (170 – 230 ms, dark gray bar, middle) and a late interval (370 – 430 ms, light gray bar, right) for comparison across conditions. Error bars indicate ± 1 SEM. The mean force responses were fit with a logistic function to the target responses of each time interval separately. E, Mean force responses to isolated cursor displacements analog to D.
The force response against the channel wall to visual perturbations has now been used in many studies (Franklin and Wolpert, 2008; Dimitriou et al., 2013; Reichenbach et al., 2014). Instead of measuring the lateral deviations in the arm trajectory, this technique quantifies the restorative force that participants apply to the handle directly and has several advantages. First, this technique has a better signal-to-noise ratio, because the trial-by-trial variability is lower for force measurements than acceleration of the arm. One of the main reasons is that force can be measured directly, whereas acceleration is normally estimated as the second derivative of position. Second, this normally leads to earlier detection of the corrective response and hence a better characterization of the gain of the fast response mechanisms. Finally, on these channel trials, the reaction of the participant does not change the trajectory of the cursor. Thus, the size of the visual error on the screen is always the same and not

Figure 2. Models and exemplary predictions. A, The difference vector model combines estimates of target and hand positions into a single difference vector, which is sent to the motor system for corrective responses. B, The weighted difference vector allows different weightings between target and hand estimate. C, The multichannel model considers two separate feedback pathways through the visual and motor systems. The force response produced by each model depends on any nonlinearity in the motor system that can be estimated using the isolated target and cursor responses. D, The theoretical output of each of the models for a simple logistic function representing the relationship between displacement size and corrective response. The colors indicate the size and direction of the corrective forces. The black dashed line indicates the diagonal in which the difference vector is — 1 cm. E, Theoretical predicted force responses for the — 1 cm difference vector diagonal for the difference vector (blue), weighted difference vector (green), and multichannel (red) model.
influenced by the strength of the early response. Thus, even the late response can be quantified directly, because the visual error signal the person sees is constant.

Each experimental block consisted of 100 randomized trials. Half of the trials were probe trials, and each perturbation condition (seven target × seven cursor displacements) occurred once plus one extra zero perturbation condition. Within each of the four sessions, subjects performed 10 blocks of trials. This resulted in 40 repetitions of each perturbation condition for each subject (80 for the zero perturbation condition) and a total of 4000 trials per participant.

Data analysis

The data were analyzed using Matlab R14. Force and kinematic time series were low-pass filtered with a fifth-order, zero-phase-lag Butterworth filter with 40 Hz cutoff and aligned to the onset of the visual perturbation (or where this would have occurred). The primary data set for analyzing the corrective motor response consisted of the force data in the channel trials due to the high sensitivity of these data. In a supplementary analysis, we also used the lateral velocity in the nonchannel trials to validate and replicate the findings in an independent data set across experimental methods. The final response profile for each condition was determined relative to the mean response in the zero-cursor/zero-target perturbation condition.

Onset time. Individual onset times of the responses to the visual perturbations were determined only for the conditions with isolated target or cursor perturbations (Fig. 1). The correction onsets were assessed based on the force data for each perturbation type (target or cursor) and perturbation level (1, 2, or 3 cm) separately. For each subject and condition, we applied t tests between the force traces of all leftward and rightward corrections until at least 20 consecutive tests (20 ms) revealed differences at a significance level of p < 0.05. The time stamp of the first of those consecutive tests was taken as the response onset time (Reichenbach et al., 2013).

Models. We tested three different models of how the corrective motor response was generated. In general, the nonlinear mapping between visual input and motor output was expressed with a logistic function (Fig. 1), where x is the size of the visual signal (dependent on the model; see below), and the parameters α and β denote the asymptote and the slope of the underlying nonlinear function. Respectively, to rigorously evaluate the model performance, we did not fit the models to the whole data set and then evaluate the goodness of fit. Instead, we built the models based only on the conditions with pure target or cursor displacements, and then evaluated how well these models predicted the conditions in which target and cursor displacements were combined. This means that for each model, we first estimated the parameters by fitting the model to the corrective response to isolated target or cursor displacements (12 conditions in total, 0/0 baseline condition subtracted). Subsequently, we predicted, based on the parameters obtained from the fits to the isolated displacements, the responses for the remaining 36 conditions of combined displacements.

The difference vector model (Fig. 2A) assumes that the force response is based on the difference between visual cursor and target displacement. In this model, the input to the nonlinear function is therefore x = (v1−v2), where v1 and v2 denote the displacements of the target and cursor, respectively. The two free parameters, α and β, were fitted to the isolated displacements and then used to predict the response to the combined displacements.

With the weighted difference vector model (Fig. 2B), we considered the possibility that the target and cursor displacement are differently weighted before the computation of their difference. For this weighting, we introduced an additional parameter, w, and thus obtained the visual input x = (1−w)v1−wv2. Values between 0 and 0.5 indicate that target information is more quickly updated than the cursor information; values between 0.5 and 1 indicate the opposite. This model therefore had three free parameters (w, α, and β).

Finally, the multichannel model postulates separate nonlinear responses to target and cursor perturbations, which then combine additively to generate the response to a combined displacement (Fig. 2C). Therefore, we fitted separate logistic function to the target displacement with x = v1, and to the cursor displacement, x = v2. This resulted in four free parameters: separate sets of parameters α and β for each visual input. For prediction of the combined displacements, the output of these two functions was simply added.

We fitted each model to the data from isolated target and cursor displacements (12 conditions) for each participant and time window independently by adjusting the parameters to minimize the residual sum of squares. Based on these parameter estimates, we calculated the model output for the trials with combined displacements (36 conditions). For model evaluation, we then calculated for each participant and time window the proportion of predicted variance as 1 minus the error sum of squares divided by the total sum of squares, using only the 36 predicted values. This generalization procedure allowed us to compare how well the models extrapolated to independent data. Therefore, we can directly compare models with different numbers of free parameters, as overfitting of more complex models on the training data automatically leads to poorer predictive performance on the independent test data (Murphy, 2012). Note that the percentage variance predicted (as opposed to the percentage variance explained in model fitting) can become negative, as the model’s prediction can be anticorrelated with the data.

Analysis of the equal difference vector conditions. The analysis illustrated in Figure 3 provides a first intuitive view on the data. We compared conditions that yielded the same visual difference vector, and thus would require the same amount of correction. The patterns within these conditions were a first indicator of whether any difference vector model is adequately describing the data. In the “zero difference vector diagonal,” we compared the conditions where target and cursor perturbations were directed in the same direction with the same amplitude, i.e., cancel each other out and do not require any corrective movement. Additionally, we considered the next two subdiagonals and superdiagonals with the difference vectors [−2, −1, 1, 2] cm.

To obtain stable values for the early (rising phase) and late (plateau phase) corrective movements, we averaged for each condition and participant all force responses within the time windows 170–230 ms and 370–430 ms, respectively.

Analysis over the time course of a reaching movement. For evaluation of the model performances, we assessed how the computation of target and cursor information evolved over the time course of the reach in time windows of 20 ms. Data were averaged over time windows from 180 ± 10 to 420 ± 10 ms after perturbation to cover the whole time interval of the corrective response. For each participant and time window, we assessed the ability of each model to predict the response to combined displacements when the parameters were fitted to the isolated displacements only. The proportion of variance predicted by models with different numbers of parameters can therefore be directly compared (see above, Analysis of the equal difference vector conditions). We then conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period. Additionally, we conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period. Additionally, we conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period. Additionally, we conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period. Additionally, we conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period. Additionally, we conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period. Additionally, we conducted two-sided paired t tests for each time window to assess whether the predicted variance with the factors model by time window over the whole computed time period.
Figure 3. Responses to conditions with equal difference vectors. **A**, Left, Conditions with ± 2 cm difference vector. Middle left, Mean force responses (solid line) and SEM (shaded region) across all subjects. Middle right, The mean (± 1 SEM) force responses over the early interval (dark gray shaded time interval, 170 – 230 ms). The horizontal lines indicate statistical differences (A, E, p < 0.005; and B–D, p < 0.0033; Bonferroni-corrected for multiple comparisons) between conditions that have been selected to demonstrate that the responses are (Figure legend continues.)
by adding two logistic functions, both having the same parameters, $\alpha = 1.7$ and $\beta = 1.4$.

**Results**

While participants performed reaching movements to a target, we unpredictable shifted the visual target and hand positions during the movement by varying amplitudes and directions (Fig. 1 B, C). Isolated target and cursor perturbations elicited rapid corrections (onset time force response, $147.4 \pm 13.1$ ms) in the direction of the displacement for target displacements and in the opposite direction for cursor displacements (Fig. 1 D, E).

There were no significant differences between the onset times to target and cursor perturbations ($F_{(1,19)} = 1.612; p = 0.218$). Importantly, the strength of the response showed a nonlinear saturation for large displacements, an effect that was especially apparent but not exclusively present during the early response intervals (Fig. 1 D, E, middle and right columns).

This nonlinearity allowed us to distinguish three models of how the nervous system may combine simultaneously occurring target and cursor displacements. The difference vector model (Fig. 2 A) proposes that the brain calculates the difference between a visual estimate of the hand and target location. This difference vector is then sent to the motor system to generate any necessary corrective response. Given the measured response nonlinearity for isolated target and cursor displacements, we can predict the response to any combination of the two factors. While showing an overall nonlinear shape, the response is identical for any combination of perturbations that results in the same difference vector (Fig. 2 D). Given that the overall goal is for the hand to reach the target, the final response should indeed resemble this calculation (Cisek et al., 1998; Gritsenko et al., 2009).

However, visual target and cursor estimates might be updated with different speeds (Brenner and Smeets, 2003). This may arise because the hand estimate is additionally informed by proprioceptive and efferent signals (van Beers et al., 1999; Sarlegna et al., 2003). Although our data showed comparable onset times across both cursor and target perturbations, it is still possible that the relative speed of updating differs between the two situations. Therefore, we also included a more general version of the difference vector model, the weighted difference vector model, which integrates the visual position of the cursor with an adjustment weight (Fig. 2 B). This model allows for stronger influence of either cursor or target estimate onto the difference vector estimate, where the weighing parameter may change over the course of the response. This model can account for motor responses to simultaneous cursor and target displacements with a zero difference vector (Brenner and Smeets, 2003).

Finally, the multichannel model (Fig. 2 C) postulates motor responses equal to the sum of the individual responses to the visual target and cursor perturbations, after they have been transformed by the response nonlinearity. It is important to note that this model could also calculate a “difference vector,” but this calculation occurs after the main nonlinearity, for instance, in the periphery at the levels of muscles or joints. This model simulates the possibility that the central nervous system estimates and corrects the initial responses to errors in the target location and hand position independently. As opposed to the two difference vector models, the multichannel model predicts a complex, nonlinear response surface (Fig. 2 D).

The differences in these predictions are most intuitive along the diagonals in which the difference vectors are matched (Fig. 2 A–C, dashed lines). The difference vector model always predicts identical responses along these lines (Fig. 2 D, E). Importantly, it also makes a strong prediction that there will be no corrective response when the difference vector is zero. The weighted difference vector predicts distinct responses for each condition. However, for any slice through the conditions (e.g., a matched difference vector diagonal), the response either increases or decreases monotonically (Fig. 2 D, E). Only the multichannel model predicts nonmonotonic trends (increases followed by decreases or vice versa) along a line with matched difference vectors (Fig. 2 D, E). The average responses along these diagonals during an early (170–230 ms) and a late (370–430 ms) postperturbation interval provide a first visual account for model comparison (Fig. 3). The significant responses in trials in which the target and cursor perturbation canceled each other (Fig. 3 C, zero difference vector) already indicate that the simple difference vector model is not appropriate for early responses. Furthermore, all other diagonals with identical difference vectors (Fig. 3, bar plots) show significant response differences, violating the predictions of the difference vector model. Importantly, the responses show nonmonotonic trends along the diagonals (Fig. 3, middle right and right columns, arrows indicate direction of significant difference between conditions), which can qualitatively only be predicted with the multichannel model.

We also illustrate the predictions for each diagonal based on the three models (Fig. 3, right two columns, line plots), which were fitted to the data from isolated cursor and target perturbations (11 conditions). For the early response, the multichannel model seems to predict the response shape quite accurately. In the late response window, the situation is less clear. Although the data conform closely to the magnitude predicted by the difference vector models, the shape of the responses continues to be well matched by the multichannel model.

These results were very stable across participants, as can be seen in individual subject responses. The early and late force responses to isolated target and cursor perturbations for an exemplar subject feature the same early nonlinearity as the group average (Fig. 4). Similarly, the force responses along the diagonals for the same subject demonstrate the same patterns as found for the group data with responses along the zero diagonal and nonmonotonic responses across all diagonals (Fig. 5).

To evaluate the predictive power of the models continuously over the temporal evolution of the response, we divided the data into 20 ms time windows and again fitted each model to isolated target and cursor displacement conditions. We then assessed the quality of the models by evaluating the predictions on the 36 left-out conditions, in which the perturbations were combined. The proportion of variance each model predicted, relative to the overall variance of the data, is plotted in Figure 6A for the force responses in the channel trials and in Figure 7A for the lateral velocity in the nonchannel trials. The proportion of variance predicted developed differentially for the three models over the time course of movement (interaction model by time, force data, $F_{(24,456)} = 9.697, p < 0.001$; velocity data, $F_{(24,456)} = 29.417, p < .001$).
During early time points, the multichannel model outperformed the two difference vector models. In detail, the multichannel model predicted a larger proportion of the variance than the difference vector model from 180 to 220 ms (force data, $t_{(19)} > 2.33$, $p < 0.04$; velocity data, $t_{(19)} > 2.07$, $p = 0.05$) and the weighted difference vector model from 180 to 200 ms (force data, $t_{(19)} > 2.11$, $p < 0.05$) and at 180 ms (velocity data, $t_{(19)} = 2.56$, $p < 0.02$). Although it seems that all models performed more poorly for this early time period than for the later period, it is important to consider that the early responses were more variable, which led to a lower maximal predictable proportion of the variance, as illustrated by the noise ceiling (Figs. 6A, 7A, black lines). For illustrative purpose, we show the average response within the early time window and the fit of each of the models to the average data for isolated cursor and target displacements (Figs. 6B, 7B). These graphs show clearly the reason for the multichannel model dominance: the two difference vector models cannot predict the more complex response surface exhibited by the actual data. Thus, it appears that the earliest part of the feedback response is not driven by a complete difference vector computation, but by a simpler approximation in which the isolated responses to target and cursor displacements are added. This pattern is reversed in the late phase of the reach (340 – 420 ms), in which both difference vector models predict the data significantly better than the multichannel model (force and velocity data, $t_{(19)} > 2.61$, $p < 0.02$). The data in the late time window present a clear diagonal pattern that is well predicted by the difference vector models (Figs. 6C, 7C). Although the multichannel model also predicts a nearly diagonal pattern for the force data (Fig. 6C), the variance explained was significantly less (Fig. 4A). Moreover the late pattern for the velocity data (Fig. 7C) demonstrates the strong differences between the model predic-

Figure 4. Responses to isolated target or cursor perturbations for a single subject. A, Left, Mean force responses (solid line) and SEM (shaded region) across all trials for one exemplar subject to isolated target displacements plotted with respect to the time of displacement onset. The mean force responses were averaged over an early interval (170 – 230 ms, dark gray bar, middle) and a late interval (370 – 430 ms, light gray bar, right) for comparison across conditions. Error bars indicate ±1 SEM. B, Mean force responses to isolated cursor displacements.
Figure 5. Responses to conditions with equal difference vectors for a single subject. A, Left, Conditions with +2 cm difference vector. Middle left, Mean force responses (solid line) and SEM (shaded region) across all trials for the same exemplar subject as Figure 4. Middle right, The mean (± SEM) force responses over the early interval (dark gray shaded time interval, 170 – 230 ms). Right, The force responses over the later interval (light gray shaded time interval, 370 – 430 ms). B, Difference vector of +1 cm. C, Difference vector of 0 cm. D, Difference vector of −1 cm. E, Difference vector of −2 cm.
Figure 6. Time course of model fits. A. Predicted variance for the 36 combined conditions for each model as a function of the time from perturbation onset (± SEM). The black line indicates the noise ceiling, i.e., the maximal predictable variance given the intrasubject variance throughout the time. Shaded regions indicate the early and late time period from the noise ceiling, i.e., the maximal predictable variance given the intrasubject variance.

Discussion

Our results reveal the temporal evolution of the integration of visual information about the hand and the target for controlling goal-directed movements. The late motor response to combined visual changes of the hand and target follows closely the prediction of difference vector models of visually guided reaching, such as the VITE (vector-integration-to-endpoint) model (Bullock et al., 1998), which states that the response is determined by the relative difference between hand and target (Sober and Sabes, 2003). However, the early response exhibits striking differences from this prediction: it is characterized by nonmonotonic changes in end-point force across different perturbation combinations with equal difference vectors, a pattern that can only be predicted by the multichannel model. This indicates that the early responses to discrepancies between expected and actual positions are characterized by independent processing of visual hand and target information, providing a rapid signaling pathway for flexible online control. The motor output only converges to the “correct response” based on a difference vector 200 ms after the first feedback correction, i.e., ~350 ms after the visual change.

Nonlinearity in visuomotor processing

Our ability to distinguish the models depends on the nonlinearity observed in the rapid visuomotor feedback response to cursor and target perturbations. If these responses were linear, all three models would provide identical predictions. With the nonlinearity, the main difference between models is that the integration of target and cursor information occurs before the nonlinearity in the difference vector and after the nonlinearity in the multichannel model. We assumed that the nonlinearity arises later in the visuomotor processing stream, likely in the translation of visual error to appropriate motor response, a process for which extensive nonlinearities have been shown (for review, see Franklin and Wolpert, 2011b). However, it is possible that part of the nonlinearity occurs earlier in visual processing. For example, the estimation of new target and cursor positions may be biased in a nonlinear fashion due to motion signals caused by the displacement. Under that assumption, the initial superior predictions from the multichannel model would still be compatible with a relatively early integration of cursor and target information.

We believe this possibility is unlikely. First, early visual processing is well understood, with clear evidence that visual signals represent accurate information about retinal position (Nakayama, 1985; Burr and Thompson, 2011). We know no evidence and 3. The p values above those quantify the comparison difference vector versus the multichannel model (two-sided t tests). B, Force responses during the early time window (170–230 ms after perturbation onset) plotted across all cursor and target displacements. Top left, Experimental data. Top right, Predicted responses of the difference vector model. Bottom left, Predicted responses from the multichannel model. Bottom right, Predicted responses from the weighted difference vector model. Note that each participant’s equal-difference vector diagonal were monotonously increasing; only their average partially displays a nonmonotonous trend. Models are fitted to each subject’s average data from isolated target and cursor perturbations, and the extrapolated predictions are shown for combined perturbations. C, Force responses during the late time window (370–430 ms after perturbation onset) and model predictions.
indicating that visual motion induces a bias in the perceived position to overestimate small displacements and underestimate large displacements. Second, in later time windows, the difference vector models outperform the multichannel model, indicating that the visuomotor system exhibits a nonlinearity after the integration of cursor and target information, i.e., in the motor response generation. This nonlinearity must also be present in the motor response generation in the early phase, forcing us to conclude that the integration of cursor and target responses occurs after the calculation of the motor response for early time points.

This is not to say that the apparent visual motion induced by the displacements has no influence on the observed nonlinearity. Previous studies indicate that both the position and motion direction of the cursor contribute to the early visuomotor feedback response (Saunders and Knill, 2004). Our results indicate that in the early phase, the responses to the cursor and target information are formed independently and integrated at a later stage, whereas in later stages of the response, the main nonlinearity occurs after the integration of the two signals.

**Why multichannel processing?**

Independent processing of visual hand and target information is clearly only an approximate solution, but may provide an efficient implementation to ensure the high speed necessary for online control (Day and Lyon, 2000; Day and Brown, 2001; Franklin and Wolpert, 2008; Reichenbach et al., 2009). At first glance, the difference vector computation appears simple. However, if one considers that the visual position and movement direction of objects are represented in population codes with bell-shaped tuning (Shmuel and Grinvald, 1996; Deneve et al., 1999; Swindale, 2000; Bosking et al., 2002), then subtraction requires an intermediate level of nonlinear basis elements (Pouget and Sejnowski, 1997; Pouget et al., 2003) and possibly recurrent activity to converge on the right solution (Makin et al., 2013; Richards et al., 2014). Therefore, this calculation may be time consuming, a luxury the visuomotor system cannot afford for time-critical online control. Processing the two visual signals independently avoids additional costs associated with integration. This temporal advance may justify a potentially inaccurate approximation that sometimes leads to unnecessary correction (compare Fig. 3C, the zero diagonal, where no response is optimal).

**Neural representation of the reach vector**

Reaching is often considered to result from the specification of the reach vector, a vector difference between the hand and the target, for planning the movement (Buneo and Andersen, 2012). While initial information regarding the location of the target is

---

**Figure 7.** Time course of model fits based on velocity data on the trials in which subjects were free to move. **A.** Predicted variance for the 36 combined conditions for each model as a function of the time from perturbation onset (±SEM). The black line indicates the noise ceiling, i.e., the maximal predictable variance given the intrasubject variance throughout the time. The proportion of variance predicted developed differentially for the three models over the time course of movement (interaction model by time, \( F_{(2,456)} = 29.417, p < .001 \)). Shaded regions indicate the early and late time periods. The \( p \) values above those quantify the comparison difference vector versus the multichannel model (two-sided \( t \) tests). **B.** Velocity responses during the early time window (170 – 230 ms after perturbation onset) plotted across all cursor and target displacements. Top left, Experimental data. Top right, Predicted responses of the difference vector model. Bottom left, Predicted responses from the multichannel model. Bottom right, Predicted responses from the weighted difference vector model. Note that each participant’s equal-difference vector diagonals were monotonously increasing; only their average partially displays a nonmonotonous trend. Models are fitted to each subject’s average data from isolated target and cursor perturbations, and the extrapolated predictions are shown for combined perturbations. **C.** Velocity responses during the late time window (370 – 430 ms after perturbation onset) and model predictions.
acquired in gaze-centered coordinates, this will be transformed into hand- or body-centered coordinates to calculate the necessary motor commands (Kalaska et al., 1997). Many studies have examined whether neurons code this information in gaze-centered reference frames or hand-centered reference frames. Extensive evidence has found neurons coding in gaze, hand, and mixed coordinate reference frames in the premotor and parietal cortices. Most relevant to our study, both neural recordings (Buneo et al., 2002; Buneo and Andersen, 2006, 2012; Pesaran et al., 2006; Bremer and Andersen, 2012) and fMRI studies (Beurze et al., 2007) found neurons tuned only for initial hand position, gaze position, or target position, even though many neurons that integrate target and hand information have also been found in a variety of areas. Importantly, the coexistence of reference frames appears to be maintained throughout the planning and movement phases (Buneo et al., 2008), suggesting that these neurons are not simply part of an initial hierarchy of reference frames to calculate a final reach vector, but that the reaching movement is simultaneously expressed in multiple coordinate reference frames. While these studies focused primarily on the delay period before the reach and did not examine reference coding during rapid perturbations, the existence of multiple maintained reference frames supports the idea that a mixture of feedback responses (independent hand and target responses as well as difference vector responses) could occur with response timing depending on factors such as computational complexity.

Neural tuning for feedforward control

We propose that the temporal evolution of spatial computations revealed here for fast feedback responses may also occur in the calculation of feedforward motor commands before movement onset. The VITE model (Bullock et al., 1998) proposes that the difference vector, i.e., required movement, is calculated in the reaches (populations in M1 are tuned to the direction and dynamics of movement), starting 60 – 80 ms after target presentation. However, even during speeded responses, muscle activity is not observed until 100 ms later (Georgopoulos et al., 1982). Indeed, the directional tuning of a large proportion of M1 neurons is locked to the onset of target presentation than to movement onset, with a significant fraction driven by both events (Rao and Donoghue, 2014). Furthermore, this study demonstrates that the directional population code in M1 evolves over time, suggesting that some sensory-to-motor transformations still must be computed to initiate movement even though visual target information is present much earlier in motor structures (Georgopoulos et al., 1989). The temporal characteristics of these neural populations exhibit a striking resemblance to the processes observed in our study. However, whereas the motor system can stall movement initiation until the computation of the difference vector has fully evolved, it utilizes these initial computations for rapid online control of movement. This suggests that the spatial computations underlying feedforward and feedback control might rely on similar processes and neuronal substrates (Desmurget and Grafton, 2000; Scott, 2004; Pruszynski et al., 2011). Thus, investigating fast feedback responses offers a behavioral window into elucidating the time course of spatial computations underlying sensorimotor control (Scott, 2004; Kurtzer et al., 2008; Resulaj et al., 2009; Franklin and Wolpert, 2011b).

The temporal evolution of the feedback response raises the question of whether a specialized neural pathway provides the fast approximate solution (a mixed model; Buneo et al., 2008), or both initial and final responses are readouts from the same recurrent network providing the best solution at each moment. The smooth transition between models in our data and the continuous evolution of directional tuning in M1 neurons (Rao and Donoghue, 2014) seem to promote the latter hypothesis, with the computations likely occurring either within M1 or further upstream in premotor and parietal areas.

References

Ashe J, Georgopoulos AP (1994) Movement parameters and neural activity in motor cortex and area 5. Cereb Cortex 4:590–600. CrossRef Medline
Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP (2005) Parietal representation of hand velocity in a copy task. J Neurophysiol 93:508–518. Medline
Batista AP, Buneo CA, Snyder LH, Andersen RA (1999) Reach plans in eye-centered coordinates. Science 285:257–260. CrossRef Medline
Beurze SM, de Lange FF, Toni I, Medendorp WP (2007) Integration of target and effector information in the human brain during reaching. J Neurophysiol 97:188–199. CrossRef Medline
Bokking WH, Crowley JC, Fitzpatrick D (2002) Spatial coding of position and orientation in primary visual cortex. Nat Neurosci 5:874–882. CrossRef Medline
Bremer LR, Andersen RA (2012) Coding of the reach vector in parietal area 5d. Neuron 75:342–351. CrossRef Medline
Bremer E, Smeets JB (2003) Fast corrections of movements with a computer mouse. Spat Vis 16:365–376. CrossRef Medline
Bullock D, Cisek P, Grossberg S (1998) Cortical networks for control of voluntary arm movements under variable force conditions. Cereb Cortex 8:48–62. CrossRef Medline
Buneo CA, Andersen RA (2006) The posterior parietal cortex: Sensorimotor interface for the planning and online control of visually guided movements. Neuropsychologia 44:2594–2606. CrossRef Medline
Buneo CA, Andersen RA (2012) Integration of target and hand position signals in the posterior parietal cortex: effects of workspace and hand vision. J Neurophysiol 108:187–199. CrossRef Medline
Buneo CA, Jarvis MR, Batista AP, Andersen RA (2002) Direct visuomotor transformations for reaching. Nature 416:632–636. CrossRef Medline
Buneo CA, Batista AP, Jarvis MR, Andersen RA (2008) Time-invariant reference frames for parietal reach activity. Exp Brain Res 188:77–89. CrossRef Medline
Burr D, Thompson P (2011) Motion psychophysics: 1985–2010. Vision Res 51:1431–1456. CrossRef Medline
Cisek P, Grossberg S, Bullock D (1998) A cortico-spinal model of reaching and proprioception under multiple task constraints. J Cogn Neurosci 10:425–444. CrossRef Medline
Day BL, Lyon IN (2000) Voluntary modification of automatic arm movements evoked by motion of a visual target. Exp Brain Res 130:159–168. CrossRef Medline
Day BL, Brown P (2001) Evidence for subcortical involvement in the visual control of human reaching. Brain 124:1832–1840. CrossRef Medline
Deneve S, Latham PE, Pouget A (1999) Reading population codes: a neural implementation of ideal observers. Nat Neurosci 2:740–745. CrossRef Medline
Desmurget M, Grafton S (2000) Forward modeling allows feedback control for fast reaching movements. Trends Cogn Sci 4:423–431. CrossRef Medline
Dimitriou M, Wolpert DM, Franklin DW (2013) The temporal evolution of feedback gains rapidly update to task demands. J Neurosci 33:10898–10909. CrossRef Medline
Duhamel IR, Bremer F, Ben Hamed S, Graf W (1997) Spatial invariance of
