

System Identification Applied to a Visuomotor Task: Near-Optimal Human Performance in a Noisy Changing Task

R. J. Baddeley,¹ H. A. Ingram,^{1,2} and R. C. Miall²

¹Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, United Kingdom, and ²University Laboratory of Physiology, Oxford OX1 3PT, United Kingdom

Sensory-motor integration has frequently been studied using a single-step change in a control variable such as prismatic lens angle and has revealed human visuomotor adaptation to often be partial and inefficient. We propose that the changes occurring in everyday life are better represented as the accumulation of many smaller perturbations contaminated by measurement noise. We have therefore tested human performance to random walk variations in the visual feedback of hand movements during a pointing task. Subjects made discrete targeted pointing movements to a visual target and received terminal feedback via a cursor the position of which was offset from the actual movement endpoint by a random walk element and a random observation element. By applying ideal observer analysis, which for this task compares human performance against that of a Kalman filter, we show that the subjects' performance was highly efficient with Fisher efficiencies reaching 73%. We then used system identification techniques to characterize the control strategy used. A "modified" delta-rule algorithm best modeled the human data, which suggests that they estimated the random walk perturbation of feedback in this task using an exponential weighting of recent errors. The time constant of the exponential weighting of the best-fitting model varied with the rate of random walk drift. Because human efficiency levels were high and did not vary greatly across three levels of observation noise, these results suggest that the algorithm the subjects used exponentially weighted recent errors with a weighting that varied with the level of drift in the task to maintain efficient performance.

Key words: visuomotor control; human; performance; movement; ideal observer analysis; system identification

Introduction

In day-to-day life, people naturally adjust the relationship between sensory inputs and motor output signals. These adaptive changes can be seen when subjects reach accurately under various manipulations of the visual input such as wearing prescription spectacles, or wedge prisms (Welsh, 1986), and by adaptation of movement trajectories to mechanical disturbances caused by clothing, by carried loads, or by experimental perturbations (Lackner and Dizio, 1994; Shadmehr and Mussa-Ivaldi, 1994; Bhushan and Shadmehr, 1999; Sainburg et al., 1999). We suggest that everyday sensory-motor changes and fluctuations are more typically attributable to an accumulation of small perturbations, rather than a single-step change, and are better represented as a random walk of the sensorimotor mapping. We propose that the adaptive process is likely to be "tuned" to respond to these random fluctuations.

To test whether people are indeed good at performing in the face of small and irregular incremental sensory perturbations, we have adopted the method of ideal observer analysis (Fisher, 1925; Burgess et al., 1981). This allows us to measure the fundamental performance of the human sensorimotor system by explicitly calculating the best possible performance that could be achieved given an optimal algorithm and then expressing subjects' perfor-

mance as a fraction of this ideal performance (Geisler, 1989). Although the exact interpretation of subjects' performance in a task depends on the particular ideal observer used in the analysis, very good performance, close to the ideal, can provide clues about how subjects are performing the task and thus yield insights into the adaptive process.

The task that we chose was one that required subjects to continually adjust their movements to cumulative small changes in the mapping between the visual and motor coordinates and one in which the nature of the ideal observer for the task is known. This was a simple pointing task (see Fig. 1) during which subjects altered their movement amplitude, on a trial-by-trial basis, to attempt to position a cursor displayed only at the end of each movement onto a visual target.

We compare the human performance in this task with that of the ideal algorithm, a Kalman filter (Kalman and Bucy, 1961) and demonstrate high Fisher efficiencies. We then modeled the observed human performance with a set of simple linear adaptive algorithms to identify which algorithm could best match the data. The two best-fitting algorithms (a finite impulse response filter and a modified delta rule) have strong similarities and suggest that the algorithm used by our subjects is based on a short-term store of previous experience, weighted by recency (Scheidt et al., 2001).

Materials and Methods

Experimental procedure

Four adult subjects (age 24–33 years; two males, two females) gave informed consent and participated in this study. The task was based on a reaching movement with their preferred hand toward visual targets to position a cursor that was offset from their true hand position onto the target.

Received Sept. 18, 2002; revised Jan. 14, 2003; accepted Jan. 17, 2003.

This research was supported by the Sussex Centre for Neuroscience Interdisciplinary Research Centre and the Wellcome Trust. H.A.I. was supported in part by a Hackett Scholarship; R.C.M. was supported by a Wellcome Senior Fellowship.

Correspondence should be addressed to Prof. R. C. Miall, University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK. E-mail: rcm@physiol.ox.ac.uk.

Copyright © 2003 Society for Neuroscience 0270-6474/03/233066-10\$15.00/0

Experiment 1. In the first experiment, subjects were denied direct vision of their hand. If there had been no perturbation, this task would simply measure the accuracy of their target pointing movements. Instead, we dynamically manipulated the mapping between the hand end-point position and the displayed cursor location (see Fig. 2A) by adding a displacement D_t on trial t using the following equations:

$$P_t = H_t + D_t, \quad (1)$$

$$D_t = W_t + N_t, \quad (2)$$

$$W_t = W_{t-1} + N(0, \sigma_w); W_0 = 0, \quad (3)$$

$$N_t = N(0, \sigma_N). \quad (4)$$

P_t is the displayed position of the cursor on trial t ; H_t is the position of the hand; D_t is the applied displacement and consisted of W_t , a random walk component (a Wiener process) with drift parameter σ_w , and a random observation noise component N_t with standard deviation σ_N . The notation $N(0, \sigma)$ represents a Gaussian distributed random variable with zero mean and SD σ . The average sum squared distance between the displayed cursor and the target position T_t was then taken as a measure of performance:

$$\text{error} = \frac{1}{N} \sum_{t=1}^N (P_t - T_t)^2. \quad (5)$$

Despite sounding abstract and difficult, this process of adjusting movements to the presence of a slowly changing displacement proved to be very natural, and our subjects had little difficulty performing it (see Figs. 2B, 3).

Visual targets and the feedback cursor were displayed via a 640×480 resolution data projector onto a horizontal back-projection screen and viewed in a semi-silvered mirror. The virtual image of target and feedback appeared in the plane of a large digitizing table, which was used to record the position of a digitizing pen held in the subject's right hand [for further details see Ingram et al. (2000)]. An opaque card placed just beneath the mirror then blocked vision of the hand. The start position was a white square target (1×1 cm) aligned with the apex of a "V" consisting of two wooden rulers, with the apex positioned at the midline and 14 cm in front of the chin rest (Fig. 1). On each trial the subject heard a computer tone and saw a white cross-shaped cursor (1×1 cm), which was aligned with the actual pen position. The subject was required to move the cursor within the starting box, and after 100 msec at this position the starting box and cursor disappeared and a target square (1×1 cm) was presented 15, 20, or 25 cm straight ahead from the starting position, in a pseudorandom order. The subject was asked to make a single comfortable movement to the target. When the pen velocity fell below a threshold of 1 cm/sec for 75 msec, a second computer tone signaled the end of the movement. The cursor was immediately displayed at a stationary position along the vector from start to final pen position, plus or minus some additional displacement distance (Eq. 1). Subjects were told that this additional displacement would alter from trial to trial in both an unpredictable way and a way that could be followed. They were instructed to "try and get the cursor to land on the target" and to minimize the error between the cursor position and the target across the session of 102 trials. Trials were self-paced but on average took 2–3 sec to complete. As the subject moved back from the target position, the cursor again disappeared, only reappearing as the pen came within 3 cm of the start position, guided by the wooden V shape. Each subject repeated eight conditions in three separate sessions, each with three different values of the SD of the observation noise ($\sigma_N = 0, 1.5, \text{ or } 3.0$ cm) and three SDs of the random walk step size ($\sigma_w = 0, 0.75, \text{ or } 1.5$ cm). The eight conditions covered all permutations of these values except the zero noise combination (i.e., $\sigma_N = \sigma_w = 0$ cm). Sixteen different arrays of 102 noise values were generated for the observation noise and for the drift noise using Matlab (The MathWorks, Inc.). They were then scaled with the appropriate SD. Each series was accepted if the resultant displacement was less than ± 10 cm, thus limiting the amplitude of desired arm movements to

the 5–35 cm range. The drift noise set and the observation noise set were independent and uncorrelated.

Experiment 2. Four subjects were then tested in the same task but with direct vision of the hand; three of these had also been tested in the no-vision condition. Removing the card beneath the semi-silvered mirror and illuminating the digitizing table below the mirror allowed vision of the hand. Thus they saw the virtual image of the target via the mirror, the static cursor displaced from the hand at the end of each trial, and also saw their own hand throughout each trial. The same exact noise sequences were used in corresponding vision and no-vision conditions for each subject; the two subjects that each completed only one of the no-vision or the direct-vision conditions were also presented with the same noise sequences.

Ideal observer analysis

Given that the observation noise component (N_t) was unpredictable, optimal performance in this task would be achieved if the subject could estimate the random walk displacement \hat{W}_t , and thus make a pointing movement of size:

$$H_t = T_t - \hat{W}_t. \quad (6)$$

Here \hat{W}_t is the subject's estimate on trial t of W_t , based on previous observations (previous errors). The optimal least-squares and maximum-likelihood algorithm for this task is the Kalman filter (Kalman and Bucy, 1961), assuming that the filter is provided with the input statistics σ_N and σ_w . The basic operation of the Kalman filter is to use the relative confidence in each new measurement (which is determined by σ_N) and in the previous estimate of \hat{W}_t (which is dependent also on σ_w) to determine how they are combined to produce the next estimate. Note that the noise statistics, σ_w and σ_N , are fixed values given to the Kalman filter; in contrast, when our subjects began each session they were not told the size of the two noise parameters. With a single known set of drift and observation noise parameters, σ_w and σ_N , it is impossible that any other observer can perform better than the Kalman filter. The Kalman filter therefore represents an upper bound on subjects' performance.

System identification by modeling subjects' performance

By modeling the movements that our subjects made in response to feedback on previous trials, we can estimate the algorithm they are using to keep their arm in calibration. We thus used a series of simple linear models to find good fits to the human data from the with-vision condition and explored whether these models failed to capture significant nonlinear components of the data.

Model 1: correct the last error. In this task the most relevant information is given by the recent history of arm movement errors, and the simplest algorithm that subjects could use is to estimate the current displacement from the last movement error:

$$\hat{W}_t = D_{t-1}. \quad (7)$$

This algorithm would be optimal to estimate the random walk component if there were no observation noise ($\sigma_N = 0$). It represents the Null hypothesis that subjects simply used the last visual error to estimate where to move to on the next trial.

Model 2: the delta rule. The delta/Rescorla-Wagner rule (Rescorla and Wagner, 1972; Sutton and Barto, 1981) creates an estimate based on previous visual errors but only makes a partial shift in the direction of each error. Thus it averages over some of the noise in the feedback. As well as being more efficient in situations with observation noise (the delta rule is the most efficient static linear estimator for this task) (Cox and Miller, 1965), it provides an excellent model of biological behavior in many different situations. Defining the error on the previous trial as $e_{t-1} = D_{t-1} - \hat{W}_{t-1}$, i.e., the difference between the actual distortion and that estimated, the delta rule estimates the distortion:

$$\hat{W}_t = \hat{W}_{t-1} + K e_{t-1}, \quad (8)$$

where K is a learning rate or weighting parameter. Note that because of the iterative nature of Equation 8, the weighting given to a particular error on trial t is an exponentially decaying function of the number of

were constrained to use the best-fitting parameters across all conditions for each subject. These were estimated using the Nelder–Mead simplex (direct search) method in Matlab (The MathWorks, Inc.) to minimize the sum-squared error. This search can get trapped in local minima, and so each search was started from eight different initial conditions and the best fit was used. Only the mixture of experts model found local minima.

Comparing models

Given that the models had different numbers of parameters, comparison between them must be independent of their complexity. More complex models, with more parameters, can provide a superior fit to a given dataset by fitting the noise in the data, even if the model is not a better description of the underlying process. So to provide a complexity-independent comparison, we used the method of N-fold, matched cross validation (Baddeley and Tripathy, 1998). The data for each subject was split into the 24 different runs (eight conditions times three repeats). Using the first 23 runs, the best-fitting (least-squares) parameters were found for a pair of models, and then the mean squared error for the excluded 24th session for both models was then calculated using these parameters. The process of calculating optimal parameters and recording the excluded error for both models was repeated 24 times for each subject, each time excluding a different session, and for all subjects.

This results in 24×5 (sessions \times subjects) pairs of numbers representing the errors for the two models. The errors were tested for a significant difference in means using a two-tailed matched-sample *t* test. The comparisons found were either very significant ($p < 10^{-5}$) or not close to significance. Any correction for multiple comparisons would therefore make no difference to our interpretation, and none was used. This procedure has reasonable characteristics as a method of non-linear model comparison (Dietterich, 1998) and has the advantage of its simplicity over Bayesian (Mackay, 1992) and information theory-derived measures (Akaike, 1974). We chose to use a nonlinear model comparison method for our linear models so that later comparison with nonlinear models is simplified.

Results

No-vision condition

Typical data from a series of 11 trials within a session are shown for a single subject in Figure 2*B*. Here the trial-by-trial errors in placing the cursor on the target are shown by vertical lines. This subject's estimation of the displacement on trial 88, for example, can be determined by the position of the pen (not the cursor) with respect to the target position on trial 89. This is indicated by the connected series of arrows. Note how the subject adjusts her estimate of the displacement over succeeding trials in response to the feedback error. Hence, on trial 88 the feedback error was large and positive, i.e., she underestimated the displacement, so on trial 89 she increased her estimate. The feedback error on this trial was smaller but still positive, and so on trial 90 her estimate increased only slightly. However, on trial 90 the feedback error was large and negative. Hence on trial 91 she decreased her estimate, and so on. Figure 3 shows a complete run of 102 trials for a typical subject: the subject's estimate of the displacement follows the actual displacement.

Regression analysis of the mean squared errors of subjects versus those of the ideal observer was first performed separately for each subject, averaging over the 102 trials per session. The slopes of these lines were then compared using an ANOVA. This revealed no significant difference among the subjects ($F_{(3,88)} = 1.52$; $p = 0.22$), and hence the subjects' data were pooled.

Because of the feedback noise (drift and observation noise), which was present and obvious to the subjects in all trials, one possible strategy was for subjects to simply move to the target position under all conditions, i.e., to ignore all feedback and not attempt to track the drifting displacement. However, the group average of mean squared errors was small in comparison with the

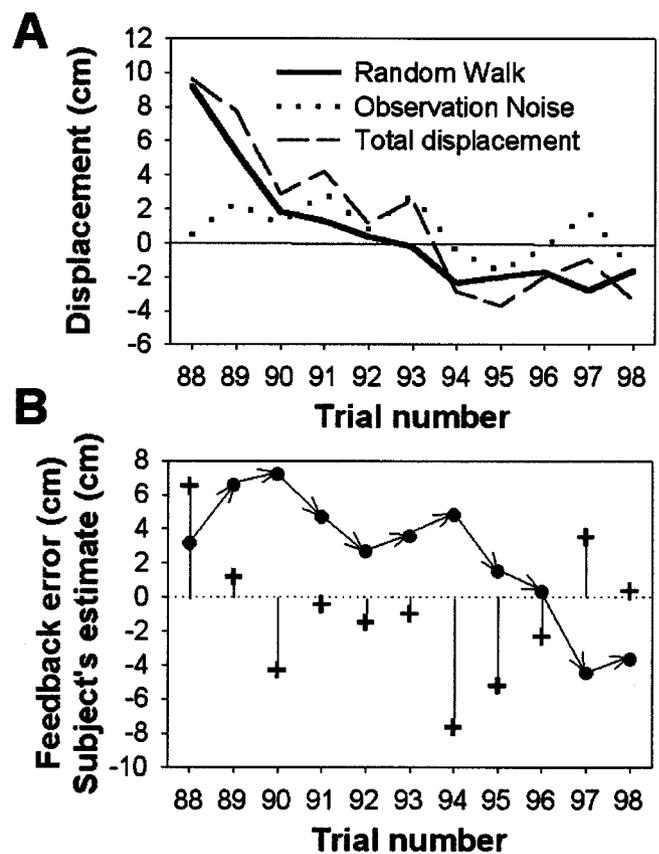


Figure 2. Typical displacement data and one subject's estimation of the displacement. *A*, Typical random walk values W_t over 11 trials are indicated by the bold solid line. Observation noise N_t (dotted line) was then added to W_t to give the total added displacement D_t (dashed line). Note how W_t is correlated with W_{t-1} , whereas N_t is uncorrelated from trial to trial. *B*, The same set of trials as in *A* showing the feedback error, $P_t - T_t$ (+) and the subject's estimate of the displacement over succeeding trials in response to the feedback error. These data are from one of the high-drift, medium-noise sessions ($\sigma_N = 1.5$ cm; $\sigma_W = 1.5$ cm).

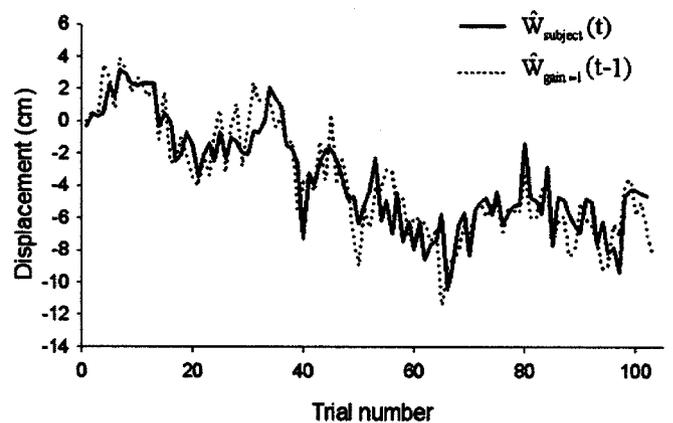


Figure 3. Typical data from a complete run of 102 trials in the high-drift, medium-noise condition ($\sigma_N = 1.5$ cm; $\sigma_W = 1.5$ cm). Here the subject's estimate (dotted lines) is shown against the actual displacement data (solid line). To emphasize the relationship between subject response and previous error, the displacement data curve has been displaced to the right by one trial. It therefore represents the performance of a system (Model 1) that estimates the displacement on trial *t* from the error observed on trial *t* - 1.

mean squared displacement, and across the six drift conditions it averaged 26% (SD 8.6%) of the mean squared feedback displacement. For example, in the highest noise condition of $\sigma_W = 1.5$ and $\sigma_N = 3$, the average mean squared displacement was 86.5

cm², whereas the group average mean square error was 19.7 cm². Hence subjects were actively and successfully tracking the drift component.

Figure 4A shows the group results of four subjects tested in eight conditions with different values of σ_W and σ_N . The data are presented as the mean Fisher efficiencies across conditions in the task. The Fisher efficiency F is the ratio between the mean square error of an ideal observer, E_I^2 (see Eq. 5), and the mean square error of the subject, E_S^2 , when the ideal observer is tested on exactly the same random walk and observation noise as the subject, i.e.:

$$F = \frac{E_I^2}{E_S^2}. \quad (16)$$

Thus subjects were able to achieve very high levels of efficiency (maximum 73%) compared with the ideal. It should be emphasized that this was not the performance of highly trained observers and that the task was quite natural, requiring little cognitive control: a conversation could be maintained at the same time that the task was being performed. The increase in Fisher efficiencies with increasing displacement noise could simply reflect the reduced ability of the ideal observer under the difficult conditions. It is true that in the limit, when $\sigma_N \gg \sigma_W$, Fisher efficiency would be high, whereas absolute performance would be low. Note, however, that performance was considerably better than expected if the subjects did not track the displacement at all. Furthermore, the Fisher efficiencies were approximately constant across the observation noise levels used, suggesting that they were not dominated by poor ideal observer performance. Further insight into the nature of the subjects' efficiency levels was gained by plotting the subjects' errors against the error of the ideal observer with the same presentation sequence (Fig. 4B). Here it is clear that there was a close relationship between ideal performance and subject performance, trial by trial. Following Burgess et al. (1981), we therefore fit a model of the form:

$$E_S^2 = M \times E_I^2 + C. \quad (17)$$

This model is equivalent to a linear regression and divides subjects' errors into two components. One is dependent on the difficulty of the task (Burgess et al., 1981), given by the slope term M , so that as the amount of noise in the task increases (i.e., as E_I^2 increases), subjects' errors also increase. The second component is independent of the task and is given by the intercept C , reflected in a constant shift of the line above that of the ideal observer. The fit of the model to the data was very good (Fig. 4B, solid line) ($r^2 = 0.811$). The slope M was found to be 1.10 [95% confidence interval (CI): 0.99–1.21] and was not significantly different from unity ($p = 0.069$; two-tailed t test). However, the intercept C was significantly different from the value of zero for

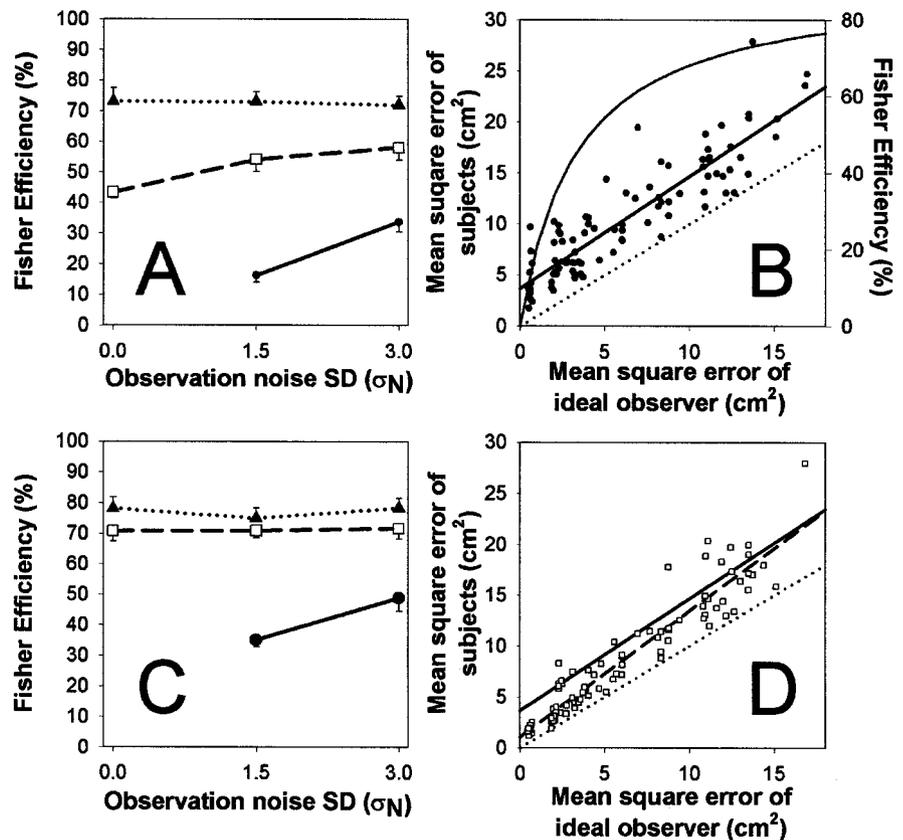


Figure 4. *A*, Subjects were very efficient (maximum 73%) at adapting to a random walk (with SD σ_W). Data are plotted as Fisher efficiencies (the ratio of the mean square error of the ideal observer to the mean square error of the subject) across all conditions (●: no drift $\sigma_W = 0$ cm; □: medium $\sigma_W = 0.75$ cm; ▲: high drift $\sigma_W = 1.5$ cm; mean efficiency ± 1 SE; $n = 12$). *B*, Subjects' run-by-run performance versus ideal performance. Subjects' mean square error is plotted against the corresponding mean square error of the ideal observer (●, $n = 96$: 4 subjects \times 8 conditions \times 3 repetitions); the **bold line** is the regression. The **dotted line** shows the performance of the ideal observer (i.e., slope equals 1 and intercept zero). The curve is an estimate of subjects' efficiency against the mean square errors of the ideal observer (Eq. 18), plotted against the axis on the right. The level of mean square error expected if subjects failed to track the drift component for $\sigma_W = 0.75$ or 1.5 would be 19.7 and 78.9 cm², respectively, in the absence of observation noise ($\sigma_N = 0$). The mean squared displacement caused only by observation noise ($\sigma_W = 0$) would be 2.0 and 8.2 cm², for $\sigma_N = 1.5$ and 3 cm, respectively. *C*, Fisher efficiency scores for subjects' performance when direct vision of the hand was allowed. The graph is in the same format as in Figure 4*A*. *D*, Run-by-run performance versus the ideal, with data from sessions with direct vision of the hand (□, $n = 96$), the regression of this data is shown by the **bold dashed line**. For comparison, the regression line from Figure 4*B* is also shown (no-vision, **solid bold line**).

the ideal observer ($C = 3.67$ cm²; 95% CI: 2.84–4.50 cm²; $p = 0.0004$).

Hence the main difference of the human data from the ideal is a constant shift (Fig. 4B), implying that most of the human inefficiency was independent of task difficulty. This inefficiency can be considered to include inaccurate arm movements, inaccurate knowledge of the position of the arm, and uncertainties in the measurement and representation of the displacement, i.e., "action noise." It is not surprising that a human should show some of these inefficiencies because the ideal observer is assumed to make accurate measurements, to store estimates with infinite precision, and to make perfectly accurate movements (an "ideal actor"). Equation 17 can be rearranged in terms of the Fisher efficiency (F ; see Eq. 16) to yield the curve shown on Figure 4B, against the right-hand axis. This curve is defined as:

$$F = \frac{E_I^2}{M \times E_I^2 + C}. \quad (18)$$

This illustrates that when there was less noise in the task, i.e., when E_I^2 was small, the effect intercept C makes the denominator

of Equation 18 much larger than the numerator, leading to a lower efficiency. When E_1^2 becomes large, the effect of the constant C becomes negligible and F approaches $1/M$. Hence the upper limit of subjects' efficiency in the no-vision task would be 1:1.10 or 91%. This ratio also indicates that if the effect of the "action noise" could be removed (or made negligible, as in the high noise limit), subjects would be able to use 91% of the available information to adjust their pointing movements. This constitutes highly efficient performance.

With-vision condition

The hypothesis that some of the subjects' inefficiency arises from inaccurate measurements or from inaccurate movements, rather than from inefficient adaptive processes, was tested by allowing subjects direct vision of their hand throughout the task. All other aspects of the task remained identical. Given direct vision, subjects should be more accurate in the control of their movements. In other words, there would be less noise associated with their production of movement H_t . They should also be able to make more accurate measurements of their movement errors (i.e., $P_t - T_t$, or by rearranging Eqs. 1 and 6, $D_t - \hat{W}_t$) and hence be able to reduce the noise in their estimate of \hat{W}_t . Therefore we would expect to see a lower intercept in the error regression analysis when vision was allowed, indicating less action noise.

The results supported this hypothesis (Fig. 4C,D). As before, there was no significant difference between the regression results of the four subjects tested in the direct vision condition ($F_{(3,88)} = 1.49$; $p = 0.22$), and again the data were also pooled across subjects.

The mean square errors were on average 20.5% (SD 5.2%) of the mean square displacement, and for the highest noise condition the group average was 17.7 cm^2 compared with the average mean squared displacement of 86.5 cm^2 . The regression of the group data obtained with direct vision of the hand to the ideal observer data was very strong (Fig. 4D, dashed line) ($r^2 = 0.896$) and had a slope of $M = 1.24$ (95% CI: 1.15–1.32; significantly greater than unity) and an intercept of $C = 1.10 \text{ cm}^2$ (95% CI: 0.48–1.72 cm^2). An ANOVA showed that the slope in the no-vision condition did not differ significantly from that in the with-vision condition ($M = 1.24$ vs 1.1; Student's t test on slopes; $t_{188} = 1.943$; $p = 0.054$). Although close to being significantly different, the 95% confidence levels in the two conditions do overlap. In contrast, the intercept with direct vision of the hand was clearly below that when vision of the hand was denied ($C = 1.1$ vs 3.67 cm^2 ; Student's t test on intercept; $t_{188} = 5.83$; $p < 0.0001$), and the 95% confidence intervals did not overlap. Hence there was a significant decrease in the level of action noise when subjects had direct vision of their hand, possibly combined with a smaller increase in error that was dependent on task difficulty.

System identification

We now compare the trial-by-trial performance from our human subject group with that of five models (see Materials and Methods). Data from one additional subject were available and are included here. The average cross-validated errors for each model are given in Table 1 with statistical significance of the improvement of fit of each model tested.

The average cross-validated error for the Null model (Model 1, "correct the last error") was 4.147 cm^2 (see example data in Fig. 3). This is the baseline against which the other more complex models are compared. The delta rule (Model 2) then provided a significantly better fit to the data with an average cross-validation

Table 1. Statistical comparison of the fit of models to human performance

	Model	Mean error of excluded sessions (cm^2)	Comparison of models
1	Correct the last error	4.147	P < 10^{-13}
2	Delta rule	2.684	
3	Kalman filter	2.683	P = 0.610
4	Finite Impulse Response	2.450	
5	Modified Delta rule	2.470	P < 10^{-8}
6	Modified Mixture of Experts	2.487	

Comparisons are made with Student's t tests; $df = 119$; $n = 5$ subjects.

error of 2.684 cm^2 ($t_{119} = 9.415$; $p < 10^{-13}$). Subjects are therefore taking more than the last error into account. Because a Kalman filter is the ideal observer in this task, we next tested its predictions as a model of their responses (Model 3). We found the best fitting constant estimates of $\hat{\sigma}_N$ and $\hat{\sigma}_W$ for each subject optimized across all eight conditions, not the true variances that varied between conditions. This fixed Kalman filter did not provide a significantly better fit than the delta rule (mean squared error of 2.683 vs 2.684 cm^2 ; $t_{119} = 0.513$; $p > 0.6$). The reason for this is that given reasonable initial estimates of the Kalman filter parameters $\hat{\sigma}_N$ and $\hat{\sigma}_W$, the value of the Kalman gain (K_t) stabilizes after only a few trials. Hence for most of the run, once K_t was stable, the Kalman filter and delta rule were equivalent, and the correlation coefficient between their predictions was $r^2 > 0.999$.

The finite impulse response model (Model 4) provided a significant improvement in fitting the human data over both the Kalman filter and the delta rule, with a mean squared cross-validated error of 2.45 cm^2 ($t_{119} = 6.120$; $p < 1.2 \times 10^{-8}$) (Table 1). Figure 5 shows the mean weighting coefficients of the optimally fitting filter from five subjects and, for comparison, the mean weightings for the best fitting delta rule model of each subject's data. The difference among the models is that the sum of all the exponentially decaying weightings of previous errors for either a delta rule or a Kalman filter is constrained to be one (Eq. 8). In contrast, for the finite impulse response model, the sum of the weighting was on average only 0.91.

The modified delta rule (Model 5) takes into account both the visual errors and some other error signal (which we assume is from proprioception) that always indicates there is no true error in that modality. Optimizing this model across the different subjects gave an average estimate of $\beta_v = 0.9$, $\beta_p = 0.1$. In other words, visual information was weighted more (90% contribution to the final prediction) than the uninformative proprioceptive information (which contributed only 10% to the final prediction). The mean squared error for this modified delta rule model was close to that of the finite impulse response model (2.48 vs 2.45 cm^2 ; $t_{119} = 1.182$; $p > 0.24$). Thus, once the effect of an inappropriate proprioceptive weighting was taken into account, an effect that pulls subjects' estimates of the displacement back toward zero, the highly constrained modified delta rule model with two free parameters provided a fit to the human data that

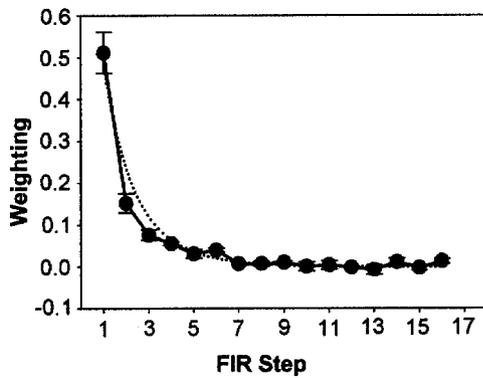


Figure 5. The average coefficients for the best 16-point finite impulse response filter fitted to each subject's data (solid line, with 1 SEM). The fine line shows the average exponential weighting found by the delta rule fits (Model 2). FIR Step, Finite impulse response coefficient number (Eq. 9).

was not significantly worse than that of the finite impulse response model, which had many more free parameters.

Testing the linearity assumption

Although the modified delta rule is the best linear approximation to the full dataset, it is important to test for significant nonlinearities. We tested for linearity in two ways. First we tested whether there was any significant structure in the residuals (ϵ_t). If the model captured everything except noise, then the difference between the best linear model (the modified delta rule) and the subjects' data would be random and uncorrelated. Hence one way to test for nonlinearity is to test whether the residuals are correlated, known as a test of whiteness (Ljung, 1999). To do this we calculated the value $\nu_{N,M}$ from the sum of the residual correlations $\hat{R}_\epsilon^N(\tau)$ separated by a spacing of τ over all N (12240) trials:

$$\hat{R}_\epsilon^N(\tau) = \frac{1}{N} \sum_{t=1}^N \epsilon_t \epsilon_{t-\tau}, \quad (19)$$

$$\nu_{N,M} = \frac{N}{\hat{R}_\epsilon^N(0)^2} \sum_{t=1}^M (\hat{R}_\epsilon^N(\tau))^2. \quad (20)$$

We summed these residual autocorrelations separated by between $M = 1 \dots 10$ time steps. The value of $\nu_{N,M}$ under the null hypothesis of no correlations should be χ^2 distributed (Ljung, 1999), but the residuals for the best-fitting modified delta rule showed highly significant residual–residual correlations ($\nu_{N,10} = 1420$; $p < 0.001$). The next sections try and place constraints on the nature of the nonlinearity.

Static (time invariant) nonlinearities

Figure 6 shows a density plot (averaged over all subjects) of the displacement D_t plotted against the inferred subjects estimate from this displacement \hat{W}_{t+1} . Also shown is the best linear fit together with the best third order polynomial. This plot shows the input–output characteristics of the system and would, if present, reveal a simple time invariant nonlinearity. There was no evidence for such a nonlinearity, with the input–output behaviour captured by a linear relationship and other terms being not significantly different from zero (cubic term $1.37 \times$

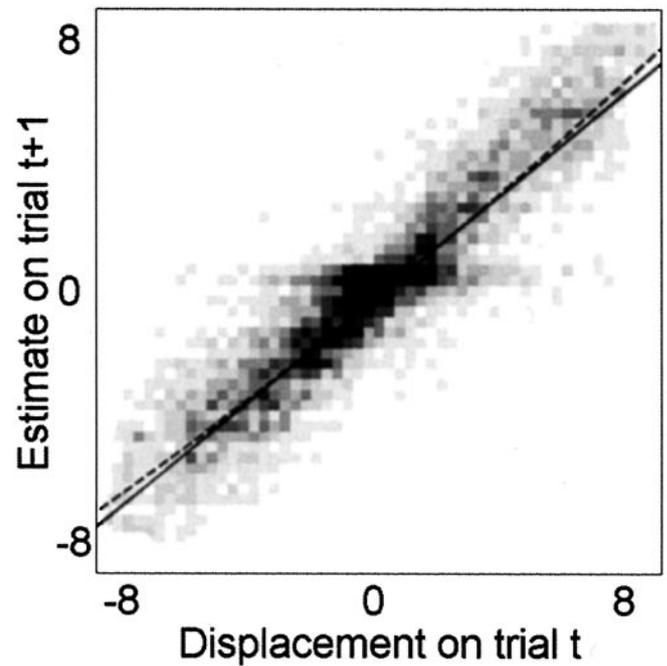


Figure 6. A density plot of the applied displacement D_{t-1} plotted against the subjects' estimate of drift \hat{W}_t . A linear regression and polynomial regression line are superimposed: for the latter, all terms higher than linear were nonsignificant.

$10^{-5} x^3$; $p = 0.79$ and the quadratic term $5.99 \times 10^{-3} x^2$; $p = 0.16$).

A second check for significant static nonlinearity is to determine whether a predictor that allowed simple nonlinearities would provide a better fit to the data than that of a linear predictor. To test this we fitted a standard mixture of experts model (Jacobs et al., 1991), in which two “experts” were each a proprioceptive modified delta rule model. By combining a mixture of two linear estimators, we can derive a nonlinear estimator of \hat{W}_t that can change its characteristics dependent on D_t . This mixture of experts showed no significant improvement in fit over the simple linear single expert (mean squared error 2.4701 cm²; comparison with single modified delta rule $t = 1.182$; $p = 0.24$). The mixture model in fact converged to a solution in which the predicted output was correlated with that of the simple linear model at the $r^2 = 0.996$ level. This analysis does not rule out the effects of complex static nonlinearities but shows that they are either not simple or not large in this data.

Temporal non-linearities

An alternative possibility is that the characteristics of the control process change over time (i.e., over condition) in an input-dependent manner. This type of nonlinearity would allow for changes in the algorithm parameters dependent on the specific task condition. To test for this we remodeled the data on a condition-by-condition basis. If there are temporal nonlinearities, then the individual conditions will be better fit by different parameters. We modeled each subject's data under each condition with the modified delta rule and tested whether there was a significant effect of condition on the best-fitting learning rate parameter K . The results are shown in Figure 7, and an ANOVA showed a significant effect of condition on best-fitting weights ($F_{(7,112)} = 40.1$; $p < 0.001$). Hence the statistics of the input have a strong influence on the processing that the subjects use.

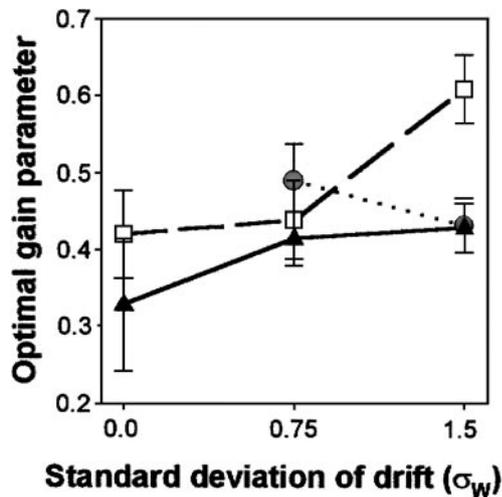


Figure 7. A plot of the learning rate parameter K (mean $n = 5$; ± 1 SD) for the best fitting modified delta rule models, calculated condition by condition for each subject. The best learning rate varied across the three levels of random walk variance tested; in comparison, the human efficiency measures were stable across observation noise levels (Fig. 4A, C).

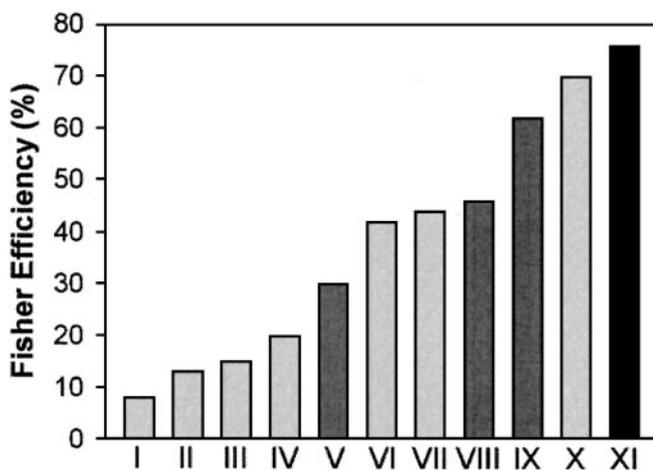


Figure 8. The highest reported human efficiencies in a number of tasks. Vision tasks are presented in light gray, auditory tasks in dark gray, and the single motor study (present results) in black. *I*, Recognizing three-dimensional (3D) objects in luminance noise (Tjan et al., 1995); *II*, 3D object classification (Liu et al., 1995); *III*, global direction of dynamic random dots (Watanianuk, 1993); *IV*, heading judgements (Crowell and Banks, 1996); *V*, detection of complex signals as a function of signal bandwidth and duration (Creelman, 1961); *VI*, letter discrimination (Parish and Sperling, 1991); *VII*, coherent visual motion (Barlow and Tripathy, 1997); *VIII*, discrimination of random, time-varying auditory spectra (Lufti, 1994); *IX*, discrimination of tonal frequency distributions (Berg, 1990); *X*, discrimination of the amplitude of a spatial sinusoid (Burgess et al., 1981); *XI*, data from the present study. The efficiency reported in the present study is higher than those reported previously.

Discussion

We have attempted to make the experimental perturbations forcing human sensorimotor recalibration more representative of the real world, and in doing so we observed high-efficiency measures relative to the ideal performance. Figure 8 shows the highest absolute efficiency values previously reported for different visual and auditory tasks in human observers, along with the highest Fisher efficiency found in the present study. As we and Burgess et

al. (1981) have shown, this absolute efficiency score includes two distinct aspects of subjects' performance in the task. First, it assesses the magnitude of the unavoidable action noise associated with performing the task, e.g., stimulus detection, response programming, controlling action execution, etc. However, we were mainly interested in the second aspect, the efficiency of the algorithm that subjects were using, and the regression analysis slope M represents this (Fig. 4B,D). It allows us to determine the fundamental efficiency limit in the task given by $1/M$. In the present study, this upper efficiency limit was 91% for the no-vision condition (81% for the with-vision condition). The study with the highest efficiency previously reported in the literature also completed a similar regression analysis (Burgess et al., 1981) and reported a limiting efficiency of 83%. Without access to the raw data of the other studies, we cannot determine the limit of their efficiency measures. However, because all of the results in Figure 8 are the maximum values from tests over a number of conditions, it seems reasonable to suggest that these maxima will at least partly reflect the limiting efficiency of each underlying algorithm. Of course, comparison of efficiencies across very different perceptual and motor tasks is problematic. Although the Fisher efficiency reflects the human performance relative to the ideal observer for each task, our comparison neglects any differences in the complexity of the algorithm used. We look forward to further use of ideal observer analysis in motor control (ideal actor analysis), so that more direct comparisons can be made.

The previously reported high-efficiency value was for the detection of two cycles of a sinusoidally varying luminance pattern superimposed on a background of noise (Burgess et al., 1981) (Fig. 8, column X). This is known to match the receptive field properties of neurons in V1, the earliest cortical area involved in vision, and Burgess et al. (1981) suggested that their high efficiency was caused by the close match between their stimulus and the V1 receptive properties. Extending this logic, the high efficiencies that we found should place constraints on any theoretical model or explanation of how subjects adjust their visuomotor mapping. Hence, there may be neural mechanisms in the motor system tuned to process and integrate movement errors over time. The predicted characteristics of such a system would be as follows: first, a neural representation of previous errors in the form of some memory trace; second, an approximately exponential weighting of these errors over time or trials to form the new estimate; and third, a mechanism to change the time or trial constant of the exponential weighting dependent on input statistics. We propose that the human motor system implements temporal filters to estimate the sensorimotor mapping as it alters throughout our lifetime.

In the current study the ideal observer had an advantage over the tested subjects because it was provided with the exact information about the drift and noise values generating each sequence. Instead, the human subjects had to estimate these parameters. However, we also show that the humans appear to adjust their adaptation process according to the specific task conditions (Fig. 7), a feature not included in our relatively simple models. So, after estimating the current noise conditions, subjects then adjust their learning rate (the weighting given to previous input errors), and we imagine that this tuning would be a continuous process. We are currently working on Bayesian algorithms to automatically estimate an appropriate learning rate for a given environment. Previous studies using perturbed visual feedback of arm movements have also demonstrated similar effects: very noisy feedback is discounted (Liu et al., 1999), and sudden sensory

perturbations may lead to incomplete adaptation because the perturbation is not credible as feedback (Kagerer et al., 1997; Ingram et al., 2000).

Thus the ideal observer's advantage (correct drift and noise values before each condition) was offset by the apparent rapid human estimation of the appropriate weighting. Of course, in the real world the amount of noise in the environment might change depending on conditions. For example, the mapping between visual input and muscle output might drift faster during fatigue or during a period of hard exercise than during a rest period. Despite this, our subjects' performance efficiency was relatively independent of the observation noise (σ_w^2), particularly at higher levels of drift (Fig. 4A). So, given their robustness to changes in observation noise rates between sessions, and their apparent rapid adjustment of the weighting parameters, we would not expect that varying the noise within an experimental session would have a large effect on our performance measures.

There have been few previous reports of such a perturbation. Most previous studies investigating visuomotor adaptation have used a single large step change as the manipulation. It is difficult to determine the appropriate ideal observer for a step change, and thus it is difficult to compute an efficiency measure comparable with those presented here. However, previous work on adaptation to angular rotation of visual feedback (Kagerer et al., 1997) or to a change in visual feedback gain (Ingram et al., 2000) has shown that subjects perform better and adapt more when the perturbation is introduced in small incremental steps than in one large change. We interpret this improvement as the result of using more natural, gradually changing stimuli. Of course, our finding that human performance exponentially weights previous errors fits well with the frequently reported exponential adaptation curves seen after step perturbations (Kagerer et al., 1997; Ingram et al., 2000). In comparison with our data (Figs. 5, 7), the exponentially decaying weighting curves would be expected to be steeper for the low noise conditions typically used by others, where averaging over trials is less advantageous, whereas these weightings would be suboptimal in the face of significant drift and observation noise.

Finally, we suggest that the slope of the regression line between ideal and human performance (Fig. 4B,D) reflects the quality of the underlying algorithm that subjects use to perform the task, once unavoidable action noise leading to a positive regression intercept is discounted. In our task, this slope was just above unity (in the with-vision task, it was significantly higher than unity, $M = 1.24$, but not significantly greater than the no-vision condition, $M = 1.1$). Both slopes suggest a small increase in errors above the ideal with increasing task difficulty. For the optimal Kalman filter algorithm, the slope would be unity; hence the subjects were performing much like the ideal algorithm for this task. Human visuomotor performance has previously been likened to Kalman filtering (Borah et al., 1988; Wolpert et al., 1995), but our system identification modeling suggests that the modified delta rule provided a significantly better fit than the Kalman filter. Further experiments will be needed to address these potential differences. The Kalman filter weights the current estimate and each new measurement in the calculation of the new estimate, according to the relative confidence in each. In the modified delta rule, similar performance is achieved by adjusting the weighting or learning rate term K (Eq. 14), and this appears to be different for each noise condition (Fig. 7). In practice, the human performance appears to be limited by what we term action noise (uncertainty about the true position of the hand with respect to the cursor and errors in reaching the exact position selected), plus

possible neural computational and execution noise, rather than by the adaptive process itself.

In summary, these results suggest that human performance in this changing sensorimotor task could be well fit by either a Kalman filter or a simple model based on the delta learning rule. These use an exponential weighting of previous movement errors, which provide noisy information about the applied perturbation, to update an estimate of the current displacement. However, a finite impulse response model achieved an even better fit, and the main difference between these models is that the sum of the FIR model weights was only 0.91, and not unity, unlike the Kalman filter or delta rule model. Thus a modified delta rule, in which a second, uninformative input acted to reduce the weight of the visual input to 0.9, achieved an equally good fit to the FIR model, with many fewer free parameters. We speculate that the second input used by the modified delta rule would be proprioception, which in this task would not provide useful information about the visual perturbation but would act to reduce the weighting given to visual errors in position (Welch and Warren, 1986; van Beers et al. 1999).

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans AC* 19:716–723.
- Baddeley RJ, Tripathy S (1998) Insights into motion perception by observer modelling. *J Opt Soc Am [A]* 15:289–296.
- Barlow HB, Tripathy SP (1997) Correspondence noise and signal pooling in the detection of coherent visual motion. *J Neurosci* 17:7954–7966.
- Berg BG (1990) Observer efficiency and weights in a multiple observation task. *J Acoust Soc Am* 88:149–158.
- Borah J, Young LR, Curry RE (1988) Optimal estimator model for human spatial orientation. *Ann NY Acad Sci* 545:51–73.
- Burgess AE, Wagner RF, Jennings RJ, Barlow HB (1981) Efficiency of human visual discrimination. *Science* 214:93–94.
- Bhushan N, Shadmehr R (1999) Computational nature of human adaptive control during learning of reaching movements in force fields. *Biol Cybern* 81:39–60.
- Cox DR, Miller HD (1965) *Theory of stochastic processes*. London: Methuen.
- Creelman CD (1961) Detection of complex signals as a function of signal bandwidth and duration. *J Acoust Soc Am* 33:89–94.
- Crowell JA, Banks MS (1996) Ideal observer for heading judgements. *Vision Res* 36:471–490.
- Dietterich TG (1998) Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput* 10:1895–1923.
- Estes WK (1950) Toward a statistical theory of learning. *Psychol Rev* 57:94–107.
- Fisher RA (1925) *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.
- Geisler WS (1989) Sequential ideal-observer analysis of visual discriminations. *Psychol Rev* 96:267–314.
- Ingram H, van Donkelaar P, Cole J, Vercher J-L, Gauthier G, Miall RC (2000) The role of proprioception and attention in a visuomotor adaptation task. *Exp Brain Res* 132:114–126.
- Jacobs RA, Jordan MI, Nowlan SJ, Hinton GE (1991) Adaptive mixture of local experts. *Neural Comput* 3:79–87.
- Kagerer FA, Contreras-Vidal JL, Stelmach GE (1997) Adaptation to gradual as compared with sudden visuo-motor distortions. *Exp Brain Res* 115:557–561.
- Kalman RE, Bucy RS (1961) New results in linear filtering and prediction. *J Basic Eng (ASME)* 83D:95–108.
- Lackner JR, Dizio P (1994) Rapid adaptation to Coriolis force perturbations of arm trajectory. *J Neurophysiol* 72:299–313.
- Liu X, Tubbesing SA, Aziz TZ, Miall RC, Stein JF (1999) Effects of visual feedback on manual tracking and action tremor in Parkinson's disease. *Exp Brain Res* 129:477–481.
- Liu Z, Knill DC, Kersten D (1995) Object classification for human and ideal observers. *Vision Res* 35:549–568.

- Ljung L (1999) *System Identification: Theory for the user*, Ed 2. Upper Saddle River, NJ: Prentice Hall.
- Lutfi RA (1994) Discrimination of random, time-varying spectra with statistical constraints. *J Acoust Soc Am* 95:1490–1500.
- Mackay D (1992) Bayesian model comparison and backprop Nets. In: *Advances in Neural Information Processing Systems*, Vol 4 (Moody JE, Hanson SJ, Lippmann RP, eds), pp 839–846. San Francisco: Morgan Kaufmann.
- Parish DH, Sperling G (1991) Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Res* 31:1399–1415.
- Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: current research and theory* (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton-Century-Crofts.
- Scheidt RA, Dingwell JB, Mussa-Ivaldi FA (2001) Learning to move amid uncertainty. *J Neurophysiol* 86:971–985.
- Shadmehr R, Mussa-Ivaldi FA (1994) Adaptive representation of dynamics during learning of a motor task. *J Neurosci* 14:3208–3224.
- Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev* 88:135–170.
- Tjan BS, Braje WL, Legge GE, Kersten D (1995) Human efficiency for recognizing 3-D objects in luminance noise. *Vision Res* 35:3053–3069.
- van Beers RJ, Sittig AC, Gon JJ (1999) Integration of proprioceptive and visual position information: an experimentally supported model. *J Neurophysiol* 81:1355–1364.
- Watamaniuk SN (1993) Ideal observer for discrimination of the global direction of dynamic random-dot stimuli. *J Opt Soc Am [A]* 10:16–28.
- Welch RB (1986) Adaptation of space perception. In: *Handbook of perception and human performance*, Vol 1 (Boff KR, Kaufman L, Thomas JP, eds), pp 24.1–24.45. New York: Wiley.
- Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor control. *Science* 269:1880–1882.