Relevance of error: what drives motor adaptation?

Kunlin Wei and Konrad Kording

Departments of Physiology, Physical Medicine and Rehabilitation, and Applied Mathematics, Northwestern University

and

Rehabilitation Institute of Chicago

345 E Superior Street

Rm O-922

60611 Chicago, IL, USA

corresponding author e-mail: k-wei@northwestern.edu

Running head: relevance of error and motor adaptation
Relevance of error: what drives motor adaptation?

Abstract

During motor adaptation the nervous system constantly uses error information to improve future movements. Today’s mainstream models simply assume that the nervous system adapts linearly and proportionally to errors. However, not all movement errors are relevant to our own action. The environment may transiently disturb the movement production, for example, a gust of wind blows the tennis ball away from its intended trajectory. Apparently the nervous system should not adapt its motor plan in the subsequent tennis strokes based on this irrelevant movement error. We hypothesize that the nervous system estimates the relevance of each observed error and only adapts strongly to relevant errors. Here we present a Bayesian treatment of this problem. The model calculates how likely an error is relevant to the motor plant and derives an ideal adaptation strategy that leads to the most precise movements. This model predicts that adaptation should be a non-linear function of the size of an error. In reaching experiments we found strong evidence for the predicted non-linear strategy. The model also explains published data on saccadic gain adaptation, adaptation to visuomotor rotations and force perturbations. Our study suggests that the nervous system constantly and effortlessly estimates the relevance of observed movement errors for successful motor adaptation.
**Introduction**

Our movements are affected by the ever-changing environment and our body. To move precisely, we need to adjust motor commands to counteract these changes using our perceived error information. Many experimental studies have addressed how people may solve this computational problem by imposing two distinct classes of perturbations onto subjects. The first class alters the visual feedback. Manipulations include wearing prism glasses to shift the view of external world (e.g., Harris 1965; van Beers et al. 1996), moving the visual target during saccades (e.g., McLaughlin 1967), and shifting or rotating the visual display of the hand in a virtual reality environment (e.g., Baddeley et al. 2003; Cheng and Sabes 2007; Elliott 1981). The second class of disturbances is to mechanically disturb movement. This includes changing the inertial property of the moving limb (e.g., Bock 1990) or introducing a novel force field by a research robot (e.g. Scheidt et al. 2001; Shadmehr and Mussa-Ivaldi 1994). Both visual and mechanical disturbances induce movement errors to which subjects readily adapt.

The role of error feedback in motor adaptation has been emphasized in many theoretical approaches. Most studies to date that focus on how the amplitude of errors influences adaptation typically assume that the size of adaptation is a linear function of past errors (Kawato et al. 1987; Scheidt et al. 2001; Thoroughman and Shadmehr 2000; Wolpert and Kawato 1998). However, we want to illustrate an important problem of this strategy with a simple example. When we return a tennis ball to a corner of the court (which challenges our opponent), we perceive how precisely the ball lands (movement error) for every serve and adjust our motor command accordingly. However, if the ball is deflected by the
opponent before landing, we should not adapt our motor command according to this error. This example demonstrates that people take the relevance of perceived errors into account during motor adaptation, i.e., they disregard the movement error that is induced by irrelevant factors. We ask how the nervous system achieves this feat and how relevance estimation influences the relationship between error and motor adaptation.

Relevance of sensory information has previously been considered in research on perception. Cue combination studies investigate how sensory inputs from different modalities are integrated into a coherent percept. If we know that two cues, for example a moving mouth and a speech signal have a common cause (i.e., both from one person) then we can combine these two signals and use our visual perception to improve speech recognition (McGurk and MacDonald 1976; Munhall et al. 1996). If we see a mouth moving without coherence to ongoing speech we will not combine the signals. It was also found that with increasing spatial or temporal disparities, multimodal cues have diminishing influence on one another (Gepshtein 2005; Gepshtein 2003; Hillis et al. 2002; Roach et al. 2006). These findings have been explained by the idea of causal inference (Ernst 2006; Knill 2003; Kording et al. 2007b), which is the term used extensively in cognitive science (Cheng and Novick 1992). It is believed that the nervous system interprets cues in terms of their causes (see also Treisman 1996; von Helmholtz 1863 (1954)). The important implication from these studies is that when cues are very different from one another in space and time, the nervous system will infer that they are not related and thus should be processed separately. Hence, the estimation of the relevance of sensory cues determines if and how the nervous system combines cues.
Here we hypothesize that motor adaptation involves an analogous process to estimate the relevance of the error information. For example, whenever the visual system detects an error, the nervous system needs to consider two hypotheses. The error may be induced by extrinsic factors, unrelated to the motor plant such as the deflected tennis ball mentioned above. In this case, the nervous system should largely disregard the visual error and only exhibit a small adaptive change, mostly driven by proprioceptive feedback. Alternatively, the error may be induced by intrinsic factors that are within the motor plant, such as fatigue of the muscles. In this case, the visual error is relevant and the motor plant should make adaptive changes large enough to overcome the error on future movements. As our perception is inherently noisy, the process of inferring the relevance of an error always happens in the presence of uncertainty. The nervous system can thus never be completely certain about the relevance of visual errors. Rather it has to estimate the probability of the above hypotheses using the multimodal sensory information and their disparities and assign weight or credit to each hypothesis. As the CNS needs to solve the relevance estimation problem in the presence of uncertainty, the appropriate approach is to use Bayesian statistics, which is the basic modeling tool in the present study.

To test our hypothesis we examine how people adapt to visual perturbations while they perform reaching movements in virtual reality. The visual error feedback is perturbed with different magnitudes and we measure the adaptation as a function of the size of the perturbation (Fig 1A and B). We develop a Bayesian model that estimates if visual feedback is relevant and should thus be used for motor adaptation or if it is perturbed and should thus be ignored (Fig 1C). The model fits the obtained data well and also explains
previously published data on adaptation of saccades (Robinson et al. 2003), adaptation to visuomotor rotation (Wei et al. 2005), and to force perturbations (Fine and Thoroughman 2007). Our results show that motor adaptation is a non-linear function of the size of the error and that this nonlinearity can be predicted by a model of relevance estimation.
Methods

Experimental setup

Seven naïve subjects participated in the study after providing informed consent. All procedures were approved by the institutional review board of Northwestern University. All subjects were right-handed with normal or corrected-to-normal vision. The subjects were seated in front of a table and used their right hand to hold a robotic manipulandum (Phantom Haptics Premium 1.5). The movement of the hand was measured by the manipulandum and visual feedback was given in real time through a mirror-projection system (projector: NEC LT170) with a display frequency of 75Hz (Fig 1A). Vision of the arm and hand was occluded as the movement was performed underneath a mirror. A black cloth was draped over the shoulder of the subject to minimize residual visual information about hand movement. Each subject performed a total of 900 trials, resulting in approximately one hour of participation.

Insert Figure 1 here.

Experimental task

The visual feedback of the hand position, the starting position and the target were presented on the projection screen (Fig 1B). The starting position was represented as a hollow circle at the middle of the workspace, aligned with the midline of the seated subject. The position of the hand was indicated by a cursor. Each trial started when the subject placed the cursor at the starting position. After the hand remained stationary at that position for 100 ms, a target, represented as a circle of 0.5 cm diameter, was
displayed 15 cm to the right of the starting position. At the same time, the circle representing the starting position disappeared and this triggered the subject to make an accurate horizontal movement to the target.

**Disturbance of the visual feedback**

As soon as the subject started moving the hand towards the target, the display of the cursor was extinguished. Feedback of the cursor’s position was only given when the cursor crossed the target in the horizontal direction (y direction in Fig 1B). This feedback was perturbed in the depth direction (x direction in Fig 1B), randomly and i.i.d. by one of 9 possible values: 0, ±1, ±2, ±4 and ±8 cm. Throughout this paper we use the word disturbance to denote perturbations in depth and we use the word deviation to denote deviations of the actual hand in the same direction. After the movement, the feedback remained visible for 100 ms at the same position. Subsequently, the subject returned the hand to the starting position for the next trial. This kind of feedback is similar to the terminal feedback utilized in other visuomotor adaptation studies (Baddeley et al. 2003; Bock 1992; Cheng and Sabes 2007; Ingram et al. 2000). The instruction was to “hit” the target as accurately as possible in a single reach without corrective movement close to the target. Before starting the experiment, subjects were familiarized with the setup by practicing without visual disturbances for 3 minutes. By applying visual disturbances of 9 different sizes to the actual hand position at the end of each reach, we manipulated the size of the visual error.

**Second experiment with smaller movement amplitude**
We are interested in factors that affect relevance estimation of movement errors during adaptation. One important factor is the inherent motor variance of the subject. Our model (see below) predicts that it is easier to estimate the relevance of the visual error if there is less motor variance, allowing the subject to be more certain that the movement error is not self-produced but rather caused by irrelevant external factors. To test this hypothesis, we manipulate the motor variance of the same reaching task by asking the subject to move a smaller distance in a second experiment. We assume that variability in the final position decreases when the movement amplitude becomes smaller (analogous to Kording et al. 2007b). All subjects who participated in the first experiment subsequently participated in the second experiment which used the same experimental setup and protocol in terms of conditions, number of trials, and visual disturbances. The only difference was that the horizontal distance between the starting position and the target was 5 cm, compared to the 15 cm of the first experiment. As a result, these smaller movements had relatively large visual disturbances.

**Ideal observer model**

We develop an ideal observer model that derives from a specification of a probabilistic generative model. This ideal observer estimates the likelihood that the visual cues are relevant for the movement production, i.e., the visual cue and the proprioceptive cue both reflect the true movement error. The results of this calculation are used to derive an estimate of the position of the hand. This estimate, in turn, is used to predict how subjects adapt. In other words we assume that subjects adapt in response to their estimate of the hand position rather than the visually perceived hand position.
Generative model

We assume that vision and proprioception provide the sensory cues in estimating the hand position but they are noisy (Fig 1C). The proprioceptively perceived position $x_{\text{prop}}$ is assumed to be drawn from a Gaussian distribution $N(x_{\text{hand}}, \sigma_{\text{prop}}^2)$, while the visually perceived position $x_{\text{vis}}$ from a similar distribution $N(x_{\text{hand}}, \sigma_{\text{vis}}^2)$ in the case of unperturbed visual feedback. $x_{\text{hand}}$ is the true hand position. $\sigma_{\text{prop}}$ and $\sigma_{\text{vis}}$ are the standard deviations of proprioception and vision, respectively. Note $x_{\text{vis}}$ and $x_{\text{prop}}$ are independent conditional on $x_{\text{hand}}$ when no visual perturbation is applied. In the case of perturbed feedback, $x_{\text{vis}}$ is assumed to be independent of the true hand position and randomly drawn from a flat distribution. Lastly, we assume that subjects do not use prior knowledge about the hand position $x_{\text{hand}}$ as such $x_{\text{hand}}$ is unbiased and comes from a flat distribution.

An ideal observer will estimate the probability that visual and proprioceptive cues both reflect the hand position. The prior probabilities of relevance and irrelevance are defined as $p(\text{relevant})$ and $p(\text{irrelevant})$, respectively. Using Bayes rule, we obtain:

$$p(\text{relevant}|x_{\text{vis}}, x_{\text{prop}}) = \frac{p(\text{relevant})p(x_{\text{vis}}, x_{\text{prop}}|\text{relevant})}{p(x_{\text{vis}}, x_{\text{prop}})}$$  \hspace{1cm} (1)$$

Given that $p(\text{relevant}) + p(\text{irrelevant}) = 1$, the denominator, which is also called partition function in Bayesian statistics and is crucial for the derivation, can be rewritten as follows:
Eq. 2 does not imply that the central nervous system tries to choose between the relevant case and the irrelevant case. Given the ambiguity of available perceptual information and inherent motor noise, our system instead estimates the probability of perceptual information being relevant. For the mathematical derivation we need to calculate the probability of stimuli assuming hypothetical knowledge of the fact that the stimulus is relevant or irrelevant (represented as conditional probability terms \( p(x_{\text{vis}}, x_{\text{prop}}|\text{relevant}) \) and \( p(x_{\text{vis}}, x_{\text{prop}}|\text{irrelevant}) \), respectively). The optimal estimate then is a weighted combination of results based on relevant or irrelevant knowledge (see eq. 2). In the relevant case the visual feedback is indicative of the position of the hand, and we calculate the joint probabilities of visual and proprioceptive perception using the flat prior over the hand position \( x_{\text{hand}} \):

\[
p(x_{\text{vis}}, x_{\text{prop}}|\text{relevant}) = \int p(x_{\text{vis}}|x_{\text{hand}})p(x_{\text{prop}}|x_{\text{hand}})dx_{\text{hand}}
\]

(3)

As this process is governed by normal distributions, we can solve it analytically. This function can be rewritten as a different Gaussian \( N(x_{\text{vis}} - x_{\text{prop}}, \sigma^2) \) (denoted as \( N(x_{\text{vis}}, \sigma^2) \) below), where

\[
\sigma^2 = \frac{\sigma_{\text{vis}}^2 \sigma_{\text{prop}}^2}{\sigma_{\text{vis}}^2 + \sigma_{\text{prop}}^2}
\]

(4)

is a function of the variances of the prior and the likelihoods.

In the irrelevant case the visual feedback is not indicative of the position of the hand,
and the probability of sensory information conditional on the visual stimulus being irrelevant $p(x_{\text{vis}}, x_{\text{prop}}|\text{irrelevant})$ does not depend on $x_{\text{vis}}$. This is because $p(x_{\text{vis}}, x_{\text{prop}}|\text{irrelevant})$ is equal to $p(x_{\text{vis}}|\text{irrelevant})p(x_{\text{prop}}|\text{irrelevant})$ as $x_{\text{vis}}$ and $x_{\text{prop}}$ are conditionally independent in the irrelevant case. $p(x_{\text{vis}}|\text{irrelevant})$ does not depend on $x_{\text{vis}}$, as we had assumed a flat distribution over the irrelevant visual feedback. $p(x_{\text{prop}}|\text{irrelevant})$ does not depend on $x_{\text{vis}}$ either as $x_{\text{vis}}$ carries no information in the irrelevant case. Thus the term $p(x_{\text{vis}}, x_{\text{prop}}|\text{irrelevant})$ is constant in $x_{\text{vis}}$. By combining with eq. 2 and eq. 3 and factoring out $p(\text{relevant})$ in the numerator and the denominator, we can thus rewrite eq. 1 as:

$$p(\text{relevant}|x_{\text{vis}}, x_{\text{prop}}) = \text{const} \frac{N(x_{\text{vis}}, \sigma^2)}{N(x_{\text{vis}}, \sigma^2) + c}$$

(5)

where $c = p(x_{\text{vis}}, x_{\text{prop}}|\text{irrelevant})p(\text{irrelevant})/p(\text{relevant})$ is a constant in $x_{\text{vis}}$ and serves as one of the free parameters in the model.

The subjects are instructed to minimize the distance to the target, and we thus assume that the cost function they use is the squared error between the estimated position and the target position (the distance in depth direction), as shown:

$$Cost = [p(\text{relevant}|x_{\text{vis}}, x_{\text{prop}})(x_{\text{relevant}} - x_{\text{tar}})^2 + p(\text{irrelevant}|x_{\text{vis}}, x_{\text{prop}})(x_{\text{irrelevant}} - x_{\text{tar}})^2]$$

(6)

where $x_{\text{tar}}$ is the target position and in our experiment $x_{\text{tar}} = 0$. As the cost function is quadratic of the estimated position, the best estimate to minimize the cost function can be calculated as the weighted mean of estimated positions from the cases of relevant and irrelevant visual feedback (Fig 3C):

$$\hat{x} = p(\text{relevant}|x_{\text{vis}}, x_{\text{prop}})\hat{x}_{\text{relevant}} + p(\text{irrelevant}|x_{\text{vis}}, x_{\text{prop}})\hat{x}_{\text{irrelevant}}$$

(7)
Where $\hat{x}_{\text{relevant}}$ and $\hat{x}_{\text{irrelevant}}$ are the optimal estimates of hand position when the nervous system is certain that the visual feedback is relevant and irrelevant, respectively. The statistically optimal way to integrate multimodal sensory cues when all the cues are relevant is assumed is to weight the cues proportionally to their inverse variances (e.g. Ernst 2002). Thus we obtain:

$$\hat{x}_{\text{relevant}} = \frac{\sigma_{\text{vis}}^2 x_{\text{vis}} + \sigma_{\text{prop}}^2 x_{\text{prop}}}{\sigma_{\text{prop}}^2 + \sigma_{\text{vis}}^2}$$ (8)

If we are only interested in expected values of the estimate of the hand position, we can drop the term that is proportional to the hand position, which is usually very close to zero and unbiased (see Fig 2A):

$$\langle \hat{x}_{\text{relevant}} \rangle = \left( \frac{\sigma_{\text{vis}}^2 x_{\text{vis}}}{\sigma_{\text{prop}}^2 + \sigma_{\text{vis}}^2} \right)$$ (9)

where $\langle \rangle$ denotes the average over all trials.

In psychophysical experiments it is never possible to directly measure the subjects’ percepts (i.e., $x_{\text{vis}}$ and $x_{\text{prop}}$). However, if we are only interested in the average estimate as a function of visual error size, eq. 7 can be rewritten by dropping the second term since $\hat{x}_{\text{irrelevant}}$ on average is zero. We can also simplify eq. 8 by recognizing it is a multiplication of $x_{\text{vis}}$ with a scaling factor $\frac{\sigma_{\text{vis}}^2}{\sigma_{\text{prop}}^2 + \sigma_{\text{vis}}^2}$. Then, by combining eq. 5, eq. 7 and eq. 9 and the fact that $x_{\text{prop}}$ is on average unbiased we obtain:

$$\hat{x}_{\text{hand}} = x_{\text{vis}} s \frac{N(x_{\text{vis}}, \sigma^2)}{N(x_{\text{vis}}, \sigma^2) + c}$$ (10)

where the free parameter $s$ characterizes the magnitude of the influence of a visual disturbance on future trials. The variable $\sigma$ characterizes the typical distance up to which
a disturbance can still be explained by the assumption of relevant visual feedback and $c$ is the free parameter and also a constant that affects the form of $p(\text{relevant}|x_{\text{vis}}, x_{\text{prop}})$ (see eq. 5).

Estimation of the hand position is not only affected by visual error at the current trial but also by those from previous trials (Fig 1D). To capture this time-varying effect, visual disturbances up to 4 trials (identified as trial lag $\Delta t = 1, 2, 3, 4$) before the current trial were assessed. The choice of choosing 4 trials was based on the finding that visual disturbances did not strongly influence the performance after 4 trials in our initial data analysis. We re-define $s$ as $s(\Delta t)$ such that it is a function of lag $\Delta t$ and captures the way the influence of a visual error changes over time. $\sigma$ and $c$ are constrained to be the same across lags as they are not assumed to be time-varying. Taken together, the estimation of hand position at the current trial is determined by combined effect from four previous trials and eq. 10 is re-written as:

$$\hat{x}_{\text{hand}} = \sum_{\Delta t=1}^{4} x_{\text{vis}, \Delta t} s(\Delta t) \frac{N(x_{\text{vis}, \Delta t}, \sigma^2)}{N(x_{\text{vis}, \Delta t}, \sigma^2) + c}$$ (11)

All subsequent analysis is done with this reduced model. We assume subjects use this estimation of hand position to guide the hand to the target in the subsequent trials. We maximized the probability of the data (the deviation of the hand) given the model to infer the parameters for each subject separately.
Results

In our experiment subjects moved their hands horizontally to a fixed target. The visual feedback of the hand position was only displayed at the end of the trial and this feedback was randomly disturbed in the depth direction by varying distances for different trials. This manipulation created a discrepancy between the visual cue and the proprioceptive cue about the actual hand position. Adaptation to a visual disturbance on one trial should lead to a compensatory deviation of the hand into the opposite direction in the following trials. The deviation of the hand position is the measure of adaptation in our task. We want to test how adaptation to visual errors is related to the size of the error and whether relevance estimation of the error can account for their relationship.

Adaptation following a visual disturbance

Deviations of the hand position are shown as a function of preceding visual perturbations for a typical subject (Fig 2A). The deviations were affected both by the preceding disturbances as well as large trial to trial noise. However, we found that positive disturbances on average led to deviations that were negative in the subsequent trial and negative disturbances on average led to deviations that were positive (Fig 2A and 2B), as expected by virtually all models of adaptation. This deviation in the opposite direction of the proceeding disturbance indicates adaptation in the direction that compensates for errors. Importantly, we found that the average size of the deviation was affected in a non-linear way by the size of the preceding disturbance. Additionally, each deviation was not only affected by the single preceding disturbance but by the history of disturbances. A
disturbance had a strong effect on the immediately subsequent trial ($\Delta t = 1$) and progressively weaker effects on later trials (Fig 2B).

Of particular interest for the current study is the relationship between the size of the disturbance and adaptation in subsequent trials. If the disturbance was zero it had no effect on subsequent trials as the deviation was not significantly different from zero ($p = 0.287$, t-test). For small disturbances between -2 cm and +2 cm we found an approximately linear influence of the size of the disturbance on subsequent trials. However, for large disturbances this influence became sub-linear.

We proceeded to analyze data from all 7 subjects (Fig 3). Similar patterns have been found across subjects: the size of the visual disturbance had an interesting non-linear influence on subsequent trials and the influence of previous disturbances decreased over time. Visually inspecting the data we found that the adaptation was linear with respect to disturbances of size from -2 cm to 2 cm and became sub-linear beyond this range. To quantify this effect we fit a linear function to the range -2 cm to 2 cm and found that the average slope for lag 1 ($\Delta t = 1$) trials was $-0.049 \pm 0.006$ (mean ± sem over subjects). We fit a linear function to the full range and found the average slope to be $-0.030 \pm 0.005$ (mean ± sem). The difference between these two slopes was significant ($p < 0.05$, paired t-test). Thus adaptation was sub-linear with respect to the error size of the full range of data. It also appeared that the nonlinear shapes for different lags were scaled versions of
each other. Overall, these findings demonstrate that for small errors the classical assumption of linear adaptation (Kawato 2002; Scheidt et al. 2001; Thoroughman and Shadmehr 2000) is well justified while it needs to be revised for large errors.

Insert Figure 3 here.

**Fitting with the relevance estimation model**

To understand how the disturbance size affects adaptation we examined the relevance estimation model explained in Introduction (details in Method). The model proposes that the nervous system estimates how likely an error has been caused by either the motor system or the virtual reality system, and uses this estimate to adapt optimally. Over the set of all subjects, the model closely fit to the behavior of the subjects (Fig 3A) and could explain $90.8 \pm 1.5\%$ of the variance. If we fit the data with a linear model (one 2-parameters linear fit for data of each lag), only $68.7 \pm 4.9\%$ of the total variance could be explained. Note the linear models for fitting data of all lags have 8 free parameters as compared to 6 parameters in our relevance estimation model. We compared the relevance estimation model to the linear model and found that the log likelihood ratio favored the relevance estimation model by $-58.8 \pm 13.4$ (This is the relative log likelihood by assuming the log likelihood for the relevance estimation mode is 0). To take the model complexity into account, we used the Bayesian information criterion (BIC, see Burnham and Anderson 2002) to compare the relevance estimation model to the linear model and found that the relevance estimation model had a BIC that was lower by $66.0 \pm 13.4$ than the linear model. We thus found that the model provided a good fit to our data and was
preferable on statistical grounds.

We proceeded to analyze what properties of the model were involved in obtaining these fits. The model parameters were inferred for individual subjects separately. Two factors, $N(x_{vis}, \sigma^2)$ and $s(\Delta t)$, are central to the way our model predicts adaptation. The first factor $N(x_{vis}, \sigma^2)$ specifies that large errors should lead to low probability of relevance. The second factor $s(\Delta t)$ specifies the way information is integrated over time. $N(x_{vis}, \sigma^2)/(N(x_{vis}, \sigma^2) + c)$ is indicative of $p(relevant|x_{vis}, x_{prod})$, the probability of relevance estimated by the subject (eq. 5). We normalized this term so that the value is 1 for 0 cm visual disturbance so we could readily compare it among subjects for plotting and statistical data analysis purposes. The normalized probability of relevance varied systematically with the size of the error (Fig 3B). As expected by the model, it decreased with increasingly larger disturbance. The second factor is the way the influence of a disturbance evolves over time. A full normative model of motor adaptation would assume how properties of the body (such as the fatigue level of the muscles) change over time (Baddeley et al. 2003; Kording et al. 2007c; Korenberg and Ghahramani 2002). Such models approximately weigh past trials in an exponential fashion such that more immediate trials have more weights. Our model allows a time varying scaling factor $s(\Delta t)$ to capture this time-varying effect. The scaling factor indeed decreased roughly exponentially with increasing trial lag $\Delta t$ (Fig 3C, fitted exponential curve with $R^2 = 0.948$ and 2 degrees of freedom). Taken together, we found that adaptation was affected by both the error size and the elapsed time from its appearance.
Second experiment

The relevance estimation model predicts that the size of error relative to motor variance is a variable that determines relevance estimation and thus the nonlinearity of motor adaptation. To test this prediction we introduced a second experiment where subjects produce smaller hand movements (5 cm) than in the first experiment (15 cm). Motor noise and variability in the final position decrease when the movement amplitude becomes smaller (Fitts 1954; Harris and Wolpert 1998). We analyzed all trials immediately following trials without a visual disturbance and found the SDs of deviations were $0.55 \pm 0.07$ cm for 15 cm movements and $0.29 \pm 0.045$ cm for 5 cm movements. We used this measure to approximate motor variance. Thus, the variance of the endpoint deviation was significantly smaller for 5 cm movements than for 15 cm movements (paired t-test, $p < 0.01$). As disturbances were larger with respect to ongoing motor errors, our model predicted that human subjects should be more likely to infer that a visual error had been caused by a visual disturbance – in particular for rather large disturbances. Only for small disturbances should subjects assume that visual errors are relevant for movement production.

In this second experiment adaptation was again a nonlinear function of the disturbance (Fig 4A, black). However, the shape of the nonlinear adaptation curve looked remarkably different and the range of linear adaptation was much smaller (compare to Fig 4A, gray). The model fits indicated that the probability of inferring a common cause was generally lower compared to the first experiment (Fig 4B). The probability drops with increasingly large visual disturbances and this drop was faster than in the first experiment, which led
to a sharper curve. To compare this probability between two experiments, we performed a $2 \times 9$ two-way ANOVA on the normalized probability of relevance. The interaction effect did not reach significant level ($p = 0.11$) and the main effect on visual disturbance was significant with $F(8,108) = 19.92$ and $p < 0.0001$. More importantly, the main effect on experiment is significant with $F(1,108) = 53.52$ and $p < 0.0001$. This implies that on average subjects had lower estimates of relevance when their movements were shorter. To quantify the range of the error size within which subjects infer that errors are relevant, we calculated the error size that leads the normalized probability drop to 50% of the maximal value. We found this spatial scale for relevance detection was $6.0 \pm 1.4$ cm and $2.6 \pm 2.1$ cm for the 15 cm and the 5 cm experiments, respectively. A pair-t test confirmed that the difference between the two experiments was significant with $p < 0.01$. Similar to the first experiment, the time varying scaling factor $s(\Delta t)$ decreased roughly exponentially with increasing lag $\Delta t$. Interestingly, this scaling factor was remarkably similar to the one from the first experiment (Fig 4C) – as was predicted by our model.

Discussion

In this paper we have introduced a relevance estimation model which predicts whether an error is caused by factors within the body or by a disturbance in the extrinsic world. Our model assumes that sensory signals and movement production are noisy and thus estimating whether an error is relevant is an inference problem that involves uncertainty.
The specific assumptions in our model include Gaussian noise in vision and proprioception, a cost function minimizing the squared error to the target and independent generative processes for visual feedback and proprioceptive feedback when error information is relevant. The first two assumptions have been widely used for sensorimotor studies. The third assumption is intuitive for our experimental setup: when the visual error is artificially perturbed in virtual reality, it is indeed from a separate process other than the body. The model estimates hand positions based on the history of error information and uses this estimate to guide subsequent movements. The model, however, does not concern about details of motor execution. The model predicts that adaptation is linear for small errors, but sub-linear for large errors. By systematically perturbing visual feedback and studying the subsequent adaptation, we have demonstrated that, indeed, there is a nonlinear relationship between the error size and the magnitude of adaptation. Our model also predicts that the adaptation to the errors of the same magnitude should be less when movement variability is reduced. This prediction has been confirmed by the results from our second experiment which only differs from the first one by a smaller movement magnitude.

Understanding how the nervous system utilizes error information has been studied extensively (e.g. Wolpert and Ghahramani 2004). Most models of motor adaptation have assumed a linear dependency between adaptation and error size (Kawato 2002; Scheidt et al. 2001; Thoroughman and Shadmehr 2000). Without theoretical modeling, some recent experimental studies have found evidence for a non-linear relationship between the size or the rate of adaptation and the size of the error (Fine and Thoroughman 2007; Robinson
et al. 2003; Wei et al. 2005). Fine and Thoroughman (2007) investigated how people adapted to a velocity-dependent force field that perturbed a straight reaching movement. The force field’s gain, which determined the strength of the perturbing force, changed randomly from trial to trial. In one of the experiment conditions, the random field perturbed the movement trajectory either to the left or to the right. Subjects counteracted this perturbation by moving in the opposite direction in the next trial. Results indicated that this adaptation showed a nonlinear scaling relationship to the perturbation amplitude, a pattern similar to our results (Fig 5A).

Wei and colleagues (2005) investigated adaptation in a reaching task in virtual reality where the visual representation of the hand was rotated 30 degrees. The reaching movement was curved under initial exposure of this visuomotor rotation and in subsequent training sessions the trajectories gradually became straight. The movement error was quantified as the deviation from the straight line. In some conditions, the visual error feedback was amplified by a gain factor of 1, 2 or 3.1. Subjects adapted to the visuomotor rotation and achieved straight-line reaching trajectory with training. However, the rate of adaptation increased with initially increased gain (gain from 1 to 2) but decreased when the gain increased further (gain from 2 to 3.1). This finding indicated that the effect of visual error on adaptation rate is a non-linear function of error size (Fig 5B).

Robinson et al (2003) studied saccadic adaptation by systematically varying the size of
visual errors introduced after the initiation of saccades. They shifted the visual target after each saccade and kept the distance (error) between the eye position and the initial target position constant for each training session. This visual error signal elicited adaptation in the amplitude of saccade to compensate the visual error. Expressed as the percentage of initially intended saccade size, the visual error had a nonlinear relationship with the adaptation gain (Fig 5C).

The important implication from these studies is that though they used different experimental paradigms such as different tasks (arm movements and saccades) and different types of perturbations (visual and mechanical disturbances) the nonlinear relationship between adaptation and error size persisted. These nonlinear relationships can be explained by our relevance estimation model. The basic form of our model was kept the same with $\sigma$ capturing the nonlinear relationship between error size and adaptation. The scaling factor $s$ was changed to one scalar (instead of a vector of 4) as these studies did not account for time-varying effect of the error and they only deal with the adaptation on a single time scale. The model can explain 99.4% and 65% of the total variance for Fine and Thoroughman’s study and Robinson et al’s study (Fine and Thoroughman 2007; Robinson et al. 2003), respectively (see grey symbols in Fig 5). No variance result can be reported for Wei et al’s study (Wei et al. 2005) as we only have average data.

The nonlinear relationship between error and adaptation might be taken as indications that human subjects are indifferent to outliers (such as very large movement errors) in
sensorimotor tasks. A previous study on human decision making found that the cost function of sensorimotor learning is insensitive to outliers (Kording and Wolpert 2004). In the task, subjects expressed their judgment about the location of a distribution “center”. By comparing possible cost functions, the authors indicated that the widely used quadratic cost function was only applicable for small errors and human subjects punished large errors much less than expected. As such, the system appeared robust to outliers. We could have fit our data with a loss function other than quadratic and obtained similar results. However, if we following this logic, we would need to assume a different loss function (albeit of the same form) for the second experiment despite that the only modification of the second experiment from the first is shortened movement amplitude.

Our relevance estimation model, instead of inferring the loss function in a post hoc fashion, simply assumes that movement variability is less with smaller movement magnitude. It thus predicts that the nervous system treats outliers as less likely to be causally related to the variable being estimated. This relevance estimation naturally leads to adaptation as we observed in the experiment.

Another study (Cohen et al. 2008) found that visual perception also followed predictions of robust statistics. Subjects were asked to judge the overall orientation of two adjacent dot clusters which had their own orientations. Results indicated that subjects weight the small dot cluster less when it was further away from the large dot cluster and a model based on robust principle component analysis accounted for this finding. However, they also found a phenomenon that could not be accounted by the model: when the density of dots in both clusters increased, the small cluster was weighted less. The simple robust
statistical model instead predicted an unchanged weighting for increasing dot density. Recognizing the limitation of robust estimator that relied on numeric residuals, the authors suggested that a psychologically valid generative model is needed. Our relevance estimation model presented here naturally explains their findings. A better segmentation of the small cluster from the principal cluster can be obtained by increasing distance between them and/or increasing the dot density of the small cluster. This leads to a smaller likelihood that the small cluster belongs to the overall dot cluster and as such its influence on perception of the overall orientation is reduced. This is exactly what has been found in Cohen et al’s study but is unexplained by their robust statistical model.

Our model is based on the notion of relevance estimation, a concept closely related to causal inference in cognitive studies. Starting from early infancy, people infer the causes of observed events (Gopnik et al. 2004; Holland 1986; Saxe et al. 2005; Tenenbaum and Griffiths 2003). Typically, interpreting an event in terms of causes is automatic and effortless. Within the Bayesian framework, such phenomena have been studied in high-level cognition (e.g., Tenenbaum et al. 2006) and sensory cue combination (Kording et al. 2007a). Here we demonstrate that in sensorimotor tasks when the nervous system continuously adapts its motor commands it solves a similar inference problem: estimate how likely it is that the error feedback is relevant to the movement production. We postulate that cognitive activities, perception and motor adaptation might share a similar mechanism that estimates the relevance of feedback.

The present study may have important implications for robotic rehabilitation, where
robots assist patients to relearn impaired movement abilities (Burgar et al. 2000; Hogan and Krebs 2004). These rehabilitation programs are usually performed in a virtual reality environment. Our study suggests that feedback should be designed such that it indicates to the nervous system that both visual and proprioceptive cues are relevant for movement production instead of being artificially manipulated in the virtual reality setting. Under this circumstance, the nervous system will attribute observed errors to the body and this will facilitate the learning process by enabling faster and stronger adaptation. The nonlinear relationship between the size of the error feedback and the adaptation also suggests that there exists a “sweet spot” in terms of the magnitude of augmented feedback where subjects may learn fastest.
ACKNOWLEDGEMENT

We want to thank Julia Trommershäuser, Sascha Serwe and Jim Patton for feedback on the manuscript.

FUNDING

This work was supported by the Falk Trust and NIH grant RO1 NS057814-01.

COMPETING INTERESTS.

The authors have declared that no competing interests exist.
Reference


Gepshtein S, Burge, J., Ernst, M.O. and Banks, M.S. The combination of vision and touch depends on spatial proximity. Journal of Vision 5, 2005.


Hillis JM, Ernst MO, Banks MS, and Landy MS. Combining sensory information:


Figure Captions

Figure 1: Experimental setup and an exemplary movement trajectory. (A). Sketch of the experimental setup. The hand movement is measured by the robot and displayed as a cursor. The projector displays the cursor, the starting and target positions. Vision of the arm is occluded. (B). Typical movement trajectories (shown as black solid line) are shown in a projection from above along with the visual error feedback received by the subject (black dot). (C). Graphical model of relevance estimation. When the error is inferred as relevant both proprioceptive and visual cues are used to estimate the hand position. When visual cues are inferred to be irrelevant to movement production, it will be disregarded and only proprioceptive cues are used. The final estimate of the hand position is a weighted average of position estimates from both cases where the weights are the probability of each case given the two types of sensory cues. (D). Illustration of time-varying effect of visual feedback in previous trials. The current estimation of hand position is influenced by visual feedbacks from 4 previous trials that are marked as $\Delta t = 1, 2, 3$ and $4$, respectively.

Figure 2: The data from a typical subject. (A). Deviations from all trials plotted as a function of the previous visual disturbances ($\Delta t = 1$). Each dot stands for a single reach and the grey error bar shows the mean and standard error (Note, it is small) for each visual disturbance type. Data points are spread in $x$ direction for better visibility. (B). Mean deviations are plotted as a function of the size of visual
disturbances for the same typical subject. Each line stands for data from trials of different lags.

Figure 3: Data and model predictions for the first experiment with 15 cm movement amplitude from all subjects. Error bars denote the standard error over subjects. (A). Mean deviations and the corresponding model predictions are plotted as a function of the size of visual disturbances for all subjects. Different subplots are for trials from different trial lags $\Delta t$. (B). The normalized probability of relevant error as a function of the size of visual disturbances. (C). The estimated scaling factor is plotted as a function of trial lags. An exponential function is fitted and shown as a dash line.

Figure 4: Data and model predictions for the second experiment with 5 cm movement amplitudes from all subjects. Error bars denote the standard errors over subjects. Data and model predictions from the first experiment are plotted in grey lines for direct comparison. (A). Mean deviations from the data and corresponding model predictions are plotted as a function of the size of visual disturbances for all subjects. Different subplots are for trials from different trial lags $\Delta t$. (B). The normalized probability of relevant error is plotted as a function of the size of visual disturbances. (C). The estimated scaling factor is plotted as a function of trial lags.

Figure 5: Empirical data from three other adaptation studies and the corresponding
predictions of the relevance estimation model. (A). The study by Wei et al. (2005): the adaptation rate in a visuomotor adaptation task is plotted as a function of visual error gain (grey bars). The black bars denote the predictions from the relevance estimation model. (B). The study by Robinson et al (2003): the adaptation gain in saccades is plotted as a function of the visual error size. The line is the prediction from the relevance estimation model. (C). The study by Fine and Thoroughman (2007): the amount of adaptation in reaching movements is plotted as a function of the gain of the viscous perturbation. The grey circles are the data from the paper and the black line is the prediction from the relevance estimation model.
Starting Position
Target Position
Visual Feedback

Visual Disturbance (cm)

Projector Mirror
Projection Screen
Table
Measure Hand Movement

A

B

C

D

Relevant Visual Error
Irrelevant Visual Error

World
CNS

\( \Delta t = 1 \)
\( \Delta t = 2 \)
\( \Delta t = 3 \)
\( \Delta t = 4 \)

Trials
Normalized P(relevant) (%) against Visual Disturbance (cm)

Scaling Factor s (a.u.) against Δt (trials)

A: Delta t = 1
B: Delta t = 2
C: Delta t = 3
D: Delta t = 4

Visual Disturbance (cm) vs. Deviations (cm)

Data vs. Model Prediction
A

Deviations (cm)

Visual Disturbance (cm)

$\Delta t = 1$

$\Delta t = 2$

$\Delta t = 3$

$\Delta t = 4$

Normalized $P(\text{relevant})$ (%)

Visual Disturbance (cm)

Data Model

5 cm

15 cm

Scaling Factor (a.u.)

$\Delta t$ (trials)

5 cm

15 cm
**A: Wei et al., 2005**

- Time Constant of Adaptation (trial)
  - Visual Error Gain
    - Data
    - Model

**B: Robinson et al., 2003**

- Percentage of Adaptation Gain (%)
  - Visual Error Size (%)
    - Gain (Ns/m)

**C: Thoroughman et al., 2007**

- Mid-movement Adaptation (cm)
  - Gain (Ns/m)