

THE EXPECTED UTILITY OF MOVEMENT

Julia Trommershäuser¹ Laurence T. Maloney² Michael S. Landy²

¹: Giessen University, Department of Psychology, Otto-Behaghel-Str. 10F,
35394 Giessen, Germany;

²: New York University, Department of Psychology and Center for Neural Science,
6 Washington Place, New York, NY 10003, USA

Revised version, October 31, 2007

In Glimcher, P. W., Camerer, C. F., Fehr, E. and Poldrack, R. A. (Eds.), *Neuroeconomics: Decision Making and the Brain*, Elsevier, in preparation.

Corresponding Author:

Dr. Julia Trommershäuser
Department of Psychology
Otto-Behaghel-Str. 10F
35394 Giessen, Germany
Email: Julia.Trommershaeuser@psychol.uni-giessen.de
Phone: +49 641 99 26108, Fax: +49 641 99 26119

Number of Words (main text): 7808

Number of Figures: 9

Number of Pages: 42

Number of References: 90

Acknowledgments: We thank Paul Glimcher for his comments on an earlier draft of this manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (Emmy-Noether-Programme; grant TR 528/1-3) and the National Institutes of Health (grant EY08266).

Our survival depends on our ability to act effectively, maximizing the chances that we achieve our movement goals. In the course of a day we make many movements, each of which can be carried out in a variety of ways. Shall I reach for that wine glass quickly or slowly? Approach from the right or left? Movement planning is a form of decision making as we choose one of many possible movement strategies to accomplish any given movement goal. It is important for us to make these “motor decisions” rapidly and well. In this chapter, we consider how movements are planned and show that a certain class of movement-planning problems is mathematically equivalent to a choice among lotteries in decision-making under risk or ambiguity (see also Trommershäuser et al., 2006a). This analogy allows us to examine movement planning from a new perspective, that of the ideal economic movement planner. It also allows us to contrast how we make decisions in two very different modalities, planning of movement and traditional economic decision making.

We review our previous work on movement planning under risk in which subjects are generally found to be very good at choosing motor strategies that come close to maximizing expected gain, a result that is in contrast with that found with paper-and-pencil decision-making tasks. We discuss the implications of these different behavioral outcomes, noting the evident differences between the sources of uncertainty and how information about uncertainty is acquired in motor and economic tasks. We finally review the evidence concerning the neural coding of probability, expected movement error and expected gain in movement under risk (see also Bossaerts, 2008, p. X-X; Glimcher, 2008, p. X-X; Newsome et al., 2008, p. X-X). We begin, however, with a brief review of previous work on how biological organisms plan movement.

Movement planning as optimization

In planning a movement, the brain has to select one of many possible movement plans or strategies. The result of executing a movement strategy is an actual trajectory and it is evidently desirable that the choice of strategy satisfy the demands of the particular task and also minimize “wear and tear” on the organism.

Typical research in the field of human motor control combines theoretical and experimental approaches. For example, a participant in a motor control experiment might perform simple reaching movements to a target, often constrained to two dimensions, i.e. along the top of a table (Fig. 1A). The recorded movement trajectories are compared to the predictions of a computational model, mimicking the conditions of the experiment. Early approaches to modeling movement planning take the form of an optimization problem in which the cost function to be minimized is biomechanical and the optimization goal is to minimize some measure of stress on the muscles and joints. These models differ primarily in the choice of the cost function.

Movement planning optimizing biomechanical constraints

Possible biomechanical cost functions include measures of joint mobility (Kaminsky and Gentile, 1986; Soechting and Lacquaniti, 1981), muscle tension changes (Dornay et al., 1996), mean squared rate of change of acceleration (Flash and Hogan, 1985), mean torque change (Uno et al., 1989), total energy expenditure (Alexander, 1997) and peak work (Soechting et al., 1995). The outcome of applying these models is typically a single, deterministic trajectory that optimizes the trade-off between the goal of the movement and the biomechanical costs for the organism. These models are successful in explaining the human ability to adapt to forces applied

during movement execution (Burdet et al., 2001; Franklin et al., 2007). Although this wide variety of cost functions have been employed, nearly all have successfully modeled reaching movements as following a nearly straight path with a bell-shaped velocity profile (Fig. 1). A combination of biomechanical constraints was demonstrated by Cuijpers et al. (2004) who showed that humans will grasp an elliptical cylinder either along its major or minor axis, resulting in a stable grip, but will tend to choose the axis closer to that used for circular cylinders (i.e., the grasp that is more comfortable).

Figure 1 about here

Compensation for noise and uncertainty

The models just described yield single deterministic trajectories that do not take into account the possible consequences of motor errors due to noise in the motor system. Consider two movement plans for a reach toward a target. The first involves maximal acceleration toward the target for half of the reach, and maximal deceleration to come to a halt at the target. The second has a smooth acceleration/deceleration profile. We could consider the biomechanical costs associated with the two movements. But we can also characterize the two movements in terms of an external criterion of success: which is more likely to hit the target? In moving from the purely internal criterion of success (minimizing biomechanical costs) to an external measure, we change the nature of the movement planning problem and its solution.

Figure 2 shows two movement plans, in this case for a 10 deg saccadic eye movement, both of which take the same amount of time to arrive, on average, at the identical target location (Harris and Wolpert, 1998). The two planned movements differ in how force is distributed along the reach. Movement 1 (Fig. 2, dashed curves) begins with an explosive burst of force, rapidly rotating the eye toward the target, and ends with an equal but opposite explosive burst as the eye

approaches the target. Such an extreme use of force is often referred to as “bang-bang” control. Movement 2 (solid curves) has a smooth acceleration profile with gradual acceleration during the first half of the reach and deceleration during the second.

Figure 2 about here

Uncertainty in the motor system originates from noisy neural control signals that lead to variable motor output. The noise is signal-dependent; larger control signals lead to larger variability in motor outcome (Harris and Wolpert, 1998). Thus, faster movements, which require larger control signals, are more variable, resulting in the well-known speed-accuracy trade-off. Modeling such signal-dependent noise, Harris and Wolpert (1998) find that the movement that minimizes positional variance at the end of the movement, subject to the two constraints of movement duration and mean end point, is Movement 2 (solid curves).

Recent experimental work concerned with the planning and execution of speeded eye and arm movements indicates that the complex sequences of neural events that underlie voluntary movements are selected so as to minimize movement error (Harris and Wolpert, 1998; Sabes and Jordan, 1997; Todorov, 2004). Note that this approach is based on the notion that the end point variability is a consequence of the “biological noise” in the motor control system and therefore unavoidable.

Following the observation that movements are corrupted by considerable motor noise and do not always follow the same deterministic trajectory, Harris and Wolpert (1998) suggested that movement trajectories are selected to minimize the variance of the final eye or arm position. They proposed that the underlying determinant of trajectory planning is the minimization of the noise in the neural control signal that activates the muscles during the execution of a motor command and in the post-movement period. In their model, the final arm or eye position is

computed as a function of a (deterministic) biomechanical expression and a noisy neural signal, where the noise increases with the magnitude of the neural signal (Fig. 2). According to the model of Harris and Wolpert (1998), the planned trajectory of the eye and arm is chosen to minimize variance in the end point of the movement. The idea behind this approach is that the variability in the final position of a saccade or pointing movement is the result of the accumulated deviations of the executed trajectory from the planned trajectory over the duration of the movement. The model managed to successfully describe horizontal saccadic eye movements, hand paths for a set of point-to-point movements (Harris and Wolpert, 1998), as well as the movement trajectories measured in an obstacle avoidance task (Hamilton and Wolpert, 2002).

In a similar approach, Sabes and Jordan (1997) studied the contribution of kinematic and dynamic properties of the arm to movement planning. Subjects moved their hands between pairs of targets, avoiding an obstacle along the path. To carry out this task rapidly while avoiding the obstacle, subjects should choose a trajectory so that the direction from the obstacle to the trajectory at its nearest approach was one for which movement variability was minimal. Given the difficulty of measuring 2-dimensional movement variability during a reach, they instead modeled the sensitivity of the arm in various directions and at various positions in the workspace. They defined three sensitivity measures (kinematic, inertial and elastic), each of which provided predictions of the point of closest approach to the obstacle that was most safe. The data were in qualitative agreement with the predictions of all three sensitivity measures, and best predicted by their inertial sensitivity model (*mobility*).

Optimization of the consequences of movement

So far, we have considered the costs of movement in terms of biomechanical costs (energy, wear and tear) and movement accuracy. Both of these criteria for movement planning may be used by humans for planning movements. We propose that these costs may be additive and that each may be “priced” in terms of utility for the movement planner. However, in most movements, there will be utilities that are a function of the movement outcome itself. A reach for a wine glass that succeeds and leads to a more rapid drink may be pleasant, but the consequences of a failure (and wine spilled on the new carpet or a broken water glass due to a collision along the way) may offset any desire to speed up the reach. Thus, we seek a way to frame the trade-off between a small probability of a collision with an obstacle and a large decrease in the chances of achieving the goal of the task. Why might a subject be more willing to risk collision with an obstacle as the reward associated with successful completion of the task is increased?

There is considerable evidence that the motor system takes its own uncertainty into account when planning movements. Consider the task of moving your hand quickly to a target. The task is more difficult for shorter movement times and for smaller and more distant targets. Subjects consistently prolonged their movement time for smaller target diameters (Fitts, 1954). Under natural reaching conditions, subjects take the uncertainty associated with their movement into account and select movement times that allow the target to be hit with constant reliability (Meyer et al., 1988).

In our own work on movement planning under risk, we make the connection between movement outcome and its consequences explicit. In our movement tasks, subjects receive monetary rewards based on the outcome of their hand movement. The central difficulty for the movement planner in these tasks is that, with speeded movement, planned movement will differ

from actual movement due to motor noise. As a result, the exact outcome of the hand movement is stochastic and the choice of a movement plan simply serves to assign probabilities to each movement outcome.

In our studies, subjects pointed rapidly at stimulus configurations consisting of a small target and nearby penalty region (Fig. 3). Hand movements that ended within the green target circle yielded a small monetary reward; those ending in the red penalty circle could result in a loss. End points in the target-penalty overlap region led to the awarding of both the reward and the penalty. A time limit was imposed, and movements that arrived after the time limit resulted in large penalties. Target size and the distance between target and penalty regions were small (< 2 cm), similar in size to the subject's movement end point variability. The movement plan that maximizes expected gain under these conditions depends on the relative position of target and penalty circle, on the loss assigned to the penalty region and on the subject's end point variability as we explain next.

Figure 3 about here

How should a subject perform this task? Clearly, the subject's visuo-motor strategy should take into account motor uncertainty and the penalty structure imposed by the task. Our model of optimal performance is built on the following assumptions:

1. When the motor system selects a visuo-motor strategy, it in effect imposes a probability density on the space of possible movement trajectories that could occur once the motor strategy is executed. This probability density is likely affected by the goal of the movement, the planned duration, the possibility of visual feedback during the movement, previous training, and intrinsic uncertainty in the motor system (e.g. Dean et al., 2007; Tassinari et al., 2006). We emphasize

that the consequences for the subject are completely mediated through this probability density and we can, for the most part, ignore the details of the actual mechanisms that produce and steer the action.

2. Whatever the penalty structure of the task, the penalty incurred by the subject depends only on the motion trajectory that *actually* occurs.
3. The subject acts so as to produce maximum expected gain (MEG) as computed from the magnitude of each possible reward and penalty and the probability of incurring it.

According to this model, the goal of movement planning is to select an optimal visuo-motor movement strategy (i.e., a movement plan) that specifies a desired movement trajectory. In this model, the optimal movement strategy is the one that maximizes expected gain. The model takes into account explicit gains associated with the possible outcomes of the movement, the mover's own task-relevant variability, and costs associated with the time limits imposed on the mover.

For the conditions of this experiment, the scene is divided into four regions (Fig. 4A): the reward-only region R_1 with gain G_1 , the overlap region R_2 with gain¹ G_2 , the penalty-only region R_3 with gain G_3 , and the background region R_4 with gain $G_4=0$. We define an optimal visuo-motor strategy S as one that maximizes the subject's expected gain

$$\Gamma(S) = \sum_{i=1}^4 G_i P(R_i|S) + G_{\text{timeout}} P(\text{timeout}|S). \quad (1)$$

¹ Here we refer to outcomes as gains denoted G_i with losses coded as negative gains. The term expected gain that we use corresponds exactly to expected value in the psychological and economic literature.

Here, $P(R_i | S)$ is the probability, given a particular choice of strategy S , of reaching region R_i before the time limit ($t = \text{timeout}$) has expired,

$$P(R_i | S) = \int_{R_i^{\text{timeout}}} P(\tau | S) d\tau, \quad (2)$$

where R_i^{timeout} is the set of trajectories τ that pass through R_i at some time after the start of the execution of the visuo-motor strategy and before the timeout. The task involves a penalty for not responding before the time limit (G_{timeout}). The probability that a visuo-motor strategy S leads to a timeout is $P(\text{timeout} | S)$.

Figure 4 about here

In our experiments, subjects win or lose points by touching the reward and penalty regions on the plane of the display before the timeout. Penalties and rewards depend only on the position of the end point in this plane, so a strategy S can be identified with the mean end point on the plane (\bar{x}, \bar{y}) that results from adopting a particular choice of strategy S that results in that mean end point.

In most of our experiments, subjects' movement variance was the same in the vertical and horizontal directions, indistinguishable from a bivariate Gaussian distribution (see Fig. 4B-C for simulated data for two aim points), and remained stable throughout the experiment (e.g. Trommershäuser et al., 2003b, 2005; Wu et al., 2006). Thus, we assume that the movement end points (x, y) are distributed according to a spatially isotropic Gaussian distribution with standard deviation σ ,

$$p(x, y | \bar{x}, \bar{y}, \sigma^2) = \frac{1}{2\pi\sigma^2} \exp\left(-\left((x - \bar{x})^2 + (y - \bar{y})^2\right)/2\sigma^2\right). \quad (3)$$

The probability of hitting region R_i is then

$$P(R_i | \bar{x}, \bar{y}, \sigma^2) = \int_{R_i} p(x, y | \bar{x}, \bar{y}, \sigma^2) dx dy. \quad (4)$$

In our experiments, the probability of a timeout is effectively constant over the limited range of relevant screen locations so, for a given end point variance σ^2 , finding an optimal movement strategy corresponds to choosing a strategy with mean aim point (\bar{x}, \bar{y}) that maximizes

$$\Gamma(\bar{x}, \bar{y}) = \sum_{i=1}^4 G_i P(R_i | \bar{x}, \bar{y}, \sigma^2). \quad (5)$$

The maximum of $\Gamma(\bar{x}, \bar{y})$ corresponds to the strategy maximizing expected gain and depends on the position and magnitude of the penalty and on the distribution of the subject's end points (Fig. 4D). When the penalty is zero, the aim point maximizing expected gain (and hence the mean end point maximizing expected gain) is the center of the target region. When the penalty is non-zero, the aim point maximizing expected gain shifts away from the penalty region and, therefore, away from the center of the target. This optimal shift is larger for greater penalties, for penalty regions closer to the target, and for larger magnitudes of motor variability.

For all conditions, we compared subjects' mean end points to those of an movement planner that maximized expected gain by taking into account its own task-relevant variability. Once we measured the task-relevant variability for each subject and for each level of perturbation, our model yielded parameter-free predictions of behavior maximizing expected gain for all experimental conditions against which subject behavior could be compared.

The subjects in our experiments chose strategies maximizing expected gain (MEG), or nearly so (Fig. 5). Efficiency was defined as the amount of money won relative to the amount of money expected for a subject who used the strategy maximizing expected gain. Subjects' efficiencies were typically above 90% (Dean et al., 2007; Trommershäuser et al., 2003a,b, 2005,

2006a). Subjects chose visuo-motor strategies that came close to maximizing gain in a wide variety of simple stimulus configurations, in good agreement with the predictions for the subject maximizing expected gain (Gepshtein et al., 2007; Stritzke and Trommershäuser, 2007; Trommershäuser et al., 2003a,b, 2005, 2006a,b).

Figure 5 about here

The experiments just described focused on spatial movement uncertainty and its consequences for behavior. Time played a role but only in the time limit imposed on completion of movements to the target area. More recent experiments focus on compensation for temporal uncertainty and, more generally, the allocation of available time. Hudson et al. (under review) carried out experiments analogous to Trommershäuser et al (2003a,b) but with subjects rewarded for making touching movements that arrived at a target within a specified time window. If the subject missed the target or arrived outside of the time window, no reward was given. In different experimental conditions the subject could also be penalized for arriving early or late as summarized in (Fig. 6A). Each of the bars is a time line and the reward window is colored green. Arriving at times colored in red incurred a penalty and arriving at times colored neither red nor green incurred no penalty and also no reward. The four reward/penalty conditions in Fig. 6A were blocked and subjects were informed about the reward/penalty structure; they saw a display similar to the time bars in Fig. 6A. The challenge for the subject was to compensate for the subject's own temporal movement uncertainty. Fig. 6B illustrates the one-dimensional computation of expected gain as a function of the temporal aim point selected by the subject. The computation is analogous to that described by Trommershäuser et al. (2003a,b), but now it is computed in one temporal dimension rather than two spatial dimensions. One difference between the spatial experiments and the temporal experiment of Hudson et al. is that temporal movement

uncertainty increases with duration of the movement. Hudson et al. found that subjects chose temporal aim points in good agreement with those predicted to maximize expected gain in each condition (Fig. 6C). Moreover, each subject compensated for the increase in timing uncertainty with movements of longer duration.

Figure 6 about here

Dean, Wu and Maloney (2007) studied how well subjects traded off speed and accuracy in attempting to hit targets whose value rapidly diminished over time. Once the target appeared, its value decreased from an initial maximum to 0 over 700-1000 ms depending on condition. The subject received the value associated with the target at the instant that it was hit. If the subject missed the target, of course, no reward was given. The challenge for the subject is to determine how to invest a scarce resource (time) by finding the compromise between increasing movement accuracy (by slowing down) and maximizing the value of the target when hit (by speeding up). Dean et al. measured subjects' performance in an initial training session and used these data to estimate each subject's probability of hitting the target as a function of movement duration. This curve is shown in Fig. 7 (solid curve) together with a dashed line of negative slope that represents the rapidly decreasing value of the target. The product of these two curves is the expected gain as a function of movement duration and the movement duration associated with maximum expected gain is marked with a circle. The experiment included four conditions that differed in how rapidly the value of the target decreased.

Figure 7 about here

Dean et al. found that eight subjects increased or decreased movement duration from condition to condition in accord with predicted MEG duration but that, overall, subjects were typically 50 ms slow in arriving at the target. This delay was not costly (almost all subjects

earned 90% or more of their predicted MEG), but it was consistent across subjects and conditions. Dean et al. conjectured that this 50 ms delay might represent a tradeoff between monetary gain and biomechanical cost as proposed by Trommershäuser et al. (2003a,b). In effect, subjects sacrificed some of their potential winnings (about \$0.001/trial) to be able to move slightly less rapidly, a tradeoff of 50 ms per “millidollar”.

Battaglia and Schrater (2007) examined human performance in a task where subjects needed time to accurately estimate the location of a target and also needed time to move to and touch the target to earn rewards. The target was indicated by a random sample of points drawn from a probability density whose centroid was the center of the otherwise invisible target (Fig. 8A). The points appeared one by one across time and the longer the subject waited to move the more points s/he would see and the more accurate the visual estimate of target location. But, increased viewing time came at the expense of a reduced time allocated to the movement. Moreover, as soon as the subject initiated movement, no further dots marking the target location would appear. There were three experimental conditions with three different probability densities differing in standard deviation (“dot scatter level”). In separate control conditions, they measured the variability of reaches constrained by visual or motor variability alone. Summing these individual variances, they could predict the tradeoff between viewing and movement time that minimized the standard deviation of the end point of movement relative to the center of the target (Fig. 8B). This tradeoff changed with dot scatter level. Their “... results suggest that the brain understands how visual and motor variability depend on time and selects viewing and movement durations to minimize consequent errors.” (p. 6989). These three temporal experiments, taken together, indicate the visual system can solve optimization problems involving allocation of a fundamental scarce resource – time.

Figure 8 about here

Movement planning and decision making

The importance of arbitrary loss functions

All the models of motor performance discussed above assume that an ideal motor system would adapt to the constraints of a given task by optimizing some measure of cost or expected gain. The three different classes of models examined differed only in the cost to be minimized or gain to be maximized: biomechanical costs, movement reliability or economic expected gain. Theories applying optimal control theory to modeling motor behavior (Harris and Wolpert, 1998; Todorov, 2004) as well as applying Bayesian decision theory to modeling adaptive motor behavior (Körding and Wolpert, 2004a), depend crucially on the choice of a fixed and arbitrary loss function. Most implementations of these theories use the mean squared error as the loss function, such that doubling an error quadruples the cost. However, deviations from this assumption have been found for large motor errors, suggesting that the motor system works as a robust estimator by imposing smaller penalties for large errors relative to small than would be expected with a mean-squared-error loss function (Körding and Wolpert, 2004b).

Bayesian inference is currently the most promising candidate for modeling adaptation of the motor system to persistent changes in the movement planner's environment. A movement strategy is a mapping from sensory input² v to a movement plan $s(v)$. The expected gain associated with the choice of strategy $s(v)$ is given by

$$EG(s) = \iiint g(\tau, w) p_T(\tau | s(v)) p_V(v | w) p_W(w) dv d\tau dw, \quad (6)$$

² We follow the convention that random variables are in upper case (e.g. X) while the corresponding specific values that those variables can take on are in lower case, e.g. $p(x)$.

where W is the random state of the world with prior distribution $p_W(w)$, V is sensory information about the state of the world with likelihood distribution $p_V(v|w)$ and T is the stochastic movement trajectory resulting from the executed movement plan $s(V)$ with distribution $p_T(\tau | s(v))$. The term $g(\tau, w)$ specifies the gain resulting from an actual trajectory τ in the actual state of the world w . This gain function can incorporate the gain associated with movement outcome (i.e., whether the actual trajectory τ accomplished the goal specified by world state w), but it can also incorporate the cost of the trajectory itself (i.e., biomechanical costs). Note that gains need not be determined by w and τ ; $g(\tau, w)$ need only represent the *expected* gain under those conditions.

Bayesian decision theory makes use of Bayes' rule, which states that one should calculate the *likelihood* of the state of the world, i.e. the probability of the sensory input given the hypothesized state of the world, $p_V(v|w)$. The likelihood is integrated with the *prior* $p_W(w)$, which reflects the subject's belief about the particular state of the world before the sensory input is received. The prior may reflect how probable it is that objects in the world have a particular size or are oriented in a particular way. By multiplying the prior and the likelihood and normalizing (scaling so that the probabilities over all possible states sum to one) we can estimate the probability of the state given the sensory input $p(w|v)$, termed the *posterior* of the state. This posterior could then become the new prior belief and could be further updated based on later sensory input (see e.g. Bays and Wolpert, 2006, for a review).

Motor adaptation in agreement with Bayesian inference has recently been demonstrated in a study in which subjects made online corrections to a reaching movement based on momentary visual feedback of hand position (Körding and Wolpert, 2004a). Visual feedback, presented midway through the movement, was displaced laterally by a distance that varied

randomly from trial to trial. The range of displacements experienced over the course of many trials formed a prior probability distribution. According to the Bayesian model, the prior distribution was combined with the feedback on a given trial to provide the best estimate of the position error. For movements in which very precise visual feedback was given, the prior distribution of displacements had little influence on the estimate of hand position. However, when the visual feedback was artificially blurred, the state estimate became increasingly biased towards the mean of the prior distribution, as predicted by a Bayesian model. In a later study, Tassinari et al. (2006) asked subjects to point at targets indicated by unreliable visual information, where targets were drawn from a prior distribution. Subjects had previously been trained on the prior distribution. Subjects displayed a similar shift from the position indicated by the visual information toward the mean of the prior distribution, although the variation of shift amounts across experimental conditions was smaller than of the predicted shifts that would maximize expected gain.

A similar result was found in a task in which subjects' movements were disturbed by force pulses of varying amplitude (Körding et al., 2004). The prior probability distribution of the strength of the force perturbation could be estimated by the subject as the distribution of forces experienced over the course of the experiment. Subjects experienced two force pulses during a single movement and were told that they were of equal strength. Thus, to resist the second force pulse, subjects had to predict its strength by combining their estimate of the first pulse with their current estimate of the mean of the prior distribution. Subjects adapted their behavior in a manner consistent with the predictions of Bayesian integration of a noisy estimate of the force of the first pulse with the prior. The contribution of the prior depended appropriately on the prior information.

Learning vs. computing

Surprisingly, subjects do not show a trend of gradually approaching maximum expected gain during these movement-planning tasks under risk. However, these effects may be explained by the specific design of the studies (e.g., Dean et al., 2007; Trommershäuser et al., 2003a,b, 2005). Before the “decision making” phase of the experiment, subjects practiced the speeded motor task extensively by simply touching green targets. During this initial training period, the experimenter monitored their motor performance until subjects’ reaction times stabilized to the time constraints of the tasks and the experimenter could measure each subject’s residual motor variability. Following training, subjects learned about the gains and losses assigned to each region and were asked to try to earn as much money as they could by hitting the green circle and trying to avoid hitting the penalty region. Subjects were not explicitly instructed to work out a motor strategy that took into account the spatial locations of reward and penalty regions and the magnitude of penalty and reward, but their highly efficient performance indicates that they did so from the first trial in which rewards and penalties were specified. To summarize, in the design of Trommershäuser et al. (2003a,b) and later work (Dean *et al.*, 2007), subjects were first trained to be “motor experts” in speeded pointing towards single targets on the screen. Only then were they confronted with a task involving tradeoffs between possible rewards and penalties. As Trommershäuser et al. (2003a,b; 2005) reported, there were no obvious trends in subjects’ aim points that would suggest that subjects were modifying their decision-making strategy in response to their experience with the decision-making task (Fig. 9A).

Figure 9 about here

To see how unusual this finding is, consider applying a hypothetical learning model that changes motor strategy only gradually in response to rewards and penalties incurred (see e.g.,

Daw and Doya, 2006; Dayan and Balleine, 2002; Sutton and Barto, 1998). In the training part of our experiments, subjects learned to aim toward the center of the green circle. After the training, and before the subject's first trial in the decision-making phase of the experiment, the subject was instructed that red circles carry penalties and green circles carry rewards. What should the subject do on the first trial of the decision-making phase of the experiment? In the absence of any reward or penalty, a learning model based on reward and penalty would predict that the subject should aim at the center of the green circle, just as in the training trials. The subject would then gradually alter motor strategy in response to the rewards and penalties incurred until the final motor strategy approximated the strategy maximizing expected gain (Fig. 9B).

However, examination of the initial trials of the decision phase of the experiment (Fig. 9A) suggests that subjects immediately changed their movement strategy from that used in training to that required to optimally trade off the probabilities of hitting the reward and penalty regions. This apparent lack of learning is of great interest in that it suggests that, while subjects certainly learned to carry out the motor task in the training phases of these experiments, and learned their own motor uncertainty, they seemed not to need further experience with the decision-making task to perform as well as they did.

Motor and perceptual decisions

Comparing subjects' behavior to the predictions of an ideal performance model allows one to explore the limits of information processing during goal-directed behavior. In some of our previous work, we have varied basic parameters of the model asking how human behavior deviates from optimal behavior maximizing expected gain once the integration of sensory, motor and reward information becomes too difficult or too time costly. Subjects do not always come close to maximizing expected gain in movement planning tasks under risk. We find that subjects

are able to optimally plan their movements, as long as full information about the stimulus configuration and the assigned rewards is provided prior to movement onset (Trommershäuser et al., 2006b). Subjects fail to select strategies that maximized expected gain in motor tasks similar to that of Trommershäuser et al. (2003ab) when there is a reward region and more than one penalty region (Wu et al., 2006), when target and penalty circles are reduced in contrast and blurred (Ma-Wyatt et al., 2007), and when rewards and penalties are awarded according to a stochastic pay-off rule (Maloney et al., 2007). Moreover, in unspedded *visual* tasks analogous to those of Trommershäuser et al. (2003ab), subjects fail to compensate for trial to trial variation in uncertainty (Landy et al., 2007). Thus, while there is a collection of speeded motor tasks under risk where performance is remarkably efficient, we cannot simply generalize these results to a broader claim that performance in any perceptual or motor task under risk would be “near-optimal”.

Movement under risk, decision making under risk

In planning movement in our movement tasks under risk our subjects are effectively choosing from among a set of many possible lotteries. To see why, consider a trial in which hits on the target and penalty yield gains of +1 and -5 cents, respectively (Fig. 4A). In executing the movement, the subject chooses a strategy S which we’ve identified with the mean end point (\bar{x}, \bar{y}) . The choice of strategy fixes the probability $P(R_i | S)$ of hitting the target region, the penalty region, the region where target and penalty overlap and the background, and hence of being awarded the gains G_i associated with each region. In the decision-making literature, this combination of event probabilities $P(R_i | S)$ and associated gains G_i is called a “lottery” $L(S)$,

$$L(S) = (P(R_1, S), G_1; P(R_2, S), G_2; P(R_3, S), G_3; P(R_4, S), G_4). \quad (7)$$

An alternative movement strategy S' corresponds to a second lottery

$$L(S) = (P(R_1, S'), G_1; P(R_2, S'), G_2; P(R_3, S'), G_3; P(R_4, S'), G_4). \quad (8)$$

As illustrated in Figure 4B-C, every mean end point results in a lottery with a corresponding expected gain, i.e., the expected number of points a subject will earn, on average, having “aimed” at (\bar{x}, \bar{y}) . However, there are many other possible mean end points and corresponding lotteries, each with its associated expected gain. By choosing among all these possible strategies, the subjects in our experiments effectively select among the many possible lotteries.

The results of our experiments indicate that subjects choose strategies maximizing expected gain, or nearly so. Efficiency was defined as the amount of money won relative to the amount expected using an optimal strategy maximizing expected gain. Subjects’ efficiencies were typically above 90% (Trommershäuser et al., 2003a,b).

In contrast to the highly efficient visual-motor strategies observed during visuo-motor tasks under risk, human decision makers in decision making under risk typically fail to maximize expected gain. Expected utility theory (Bernoulli 1738/1954; von Neumann and Morgenstern, 1944) is based on the assumption that subjects assign numerical utilities to outcomes and maximize expected utility. An evident advantage of the utility hypothesis is that a wide range of consequences (e.g. biomechanical costs and money) can be measured in identical units and it becomes meaningful to seek optimal tradeoffs among them. The model we presented assumes that this is the case.

When outcomes are specified as monetary rewards (as in our experiments), utility can be a nonlinear function of monetary rewards and can also depend on subjects’ current “wealth” and consequently a subject maximizing expected utility would fail to maximize expected gain. Bernoulli (1738/1954) originally suggested that this utility function is a concave function of

value which increases quickly and then flattens out for larger values and that the shape of the utility function could explain observed risk aversion in human decision making. Non-linear utility functions applied to gains and losses have been employed in describing human performance in a variety of economic decision tasks (Bell et al., 1988; Kahneman et al., 1982; Kahneman and Tversky, 2000). There is nothing inherently suboptimal or erroneous in maximizing expected utility, however it is defined, since by definition utility is what the organism seeks to maximize.

Moreover, several factors may have contributed to subjects' tendency to maximize expected gain found in our studies (Dean et al., 2007; Trommershäuser et al., 2003a,b, 2005). In the motor task, the subject makes a long series of choices and over the course of the experiment the accumulated winnings increase. On the other hand, subjects in economic decision-making experiments typically make a single "one-shot" choice between a discrete set of lotteries. Indeed, when economic decision makers are faced with a series of decisions they tend to move closer to maximum expected gain (Redelmeier and Tversky, 1992; Wakker et al., 1997; "the house money effect:" Thaler and Johnson, 1990). Studies of risky choice find that subjects are closer to maximizing expected gain for small stakes (Camerer, 1992; Holt and Laury, 2002) and when subjects receive feedback over the course of the experiment or have prior experience with the task (Hertwig et al., 2004). All of these factors would tend to "linearize" subjects' utility functions and to move them toward maximizing expected gain.

However, human performance in decision-making tasks is markedly suboptimal by other relevant criteria. Deviations from expected utility theory include a tendency to change one's decision based on whether the lottery was described in terms of losses or gains due to an exaggerated aversion to losses (Kahneman and Tversky, 1979) and a tendency to exaggerate

small probabilities (Allais, 1953; Attneave, 1953; Lichtenstein et al., 1978; Tversky and Kahneman, 1992). This exaggeration of the frequency of low-frequency events is observed in many but not all decision-making studies (Sedlmeier et al., 1998). These distortions of probability would, if present in movement planning, would be particularly problematic. The strategies maximizing expected gain in many of the motor tasks above involve small probabilities of large losses (Figure 4) and exaggeration aversion of losses and overweighting of small probabilities would likely impair performance.

The contrast between success in “movement planning under risk” and decision making under risk is heightened by the realization that, in cognitive decision making under risk, subjects are told the exact probabilities of outcomes and thus have perfect knowledge of how their choice of strategy changes the probability of attaining each outcome. The knowledge of probabilities in equivalent motor tasks is never communicated explicitly, but acquired across a few hundred training trials and thus can equal but never exceed the knowledge available under cognitive decision making under risk. The results of our experiments imply that subjects are able to learn their own motor uncertainties very well (Gepshtein et al., 2007; Trommershäuser et al., 2005; see also Baddeley et al., 2003). These results suggest that humans are able to estimate the uncertainties associated with sensory and motor noise and make use of this knowledge to improve their performance.

In summary, the results of our work indicate that movement planning shares the same formal structure as perceptual decision making and economic decision making. Subjects in movement tasks are generally found to be very good at choosing motor strategies that come close to maximizing expected gain. In contrast, subjects in economic decision-making and perceptual estimation tasks typically fail to maximize expected gain. Moreover, the sources of uncertainty

in motor tasks are endogenous: they reflect the organism's own uncertainty in planning and executing movement while, in contrast, uncertainty in economic tasks is typically imposed by the experimenter. Thus, probabilistic information from cognition, perception, and movement has different origins and it would be of interest for future work to compare the neural circuits underlying the representation of probability in movement and economic decision making.

Neural correlates of motor and cognitive decisions

We finally summarize recent experimental work directed at understanding the neural coding of motor and cognitive decisions. Most of the current evidence results from electrophysiological recordings in monkeys measuring single-cell activity during binary-decision tasks in response to manipulations of reward and reward uncertainty (see e.g., Sugrue et al., (2005) for a review).

Following Herrnstein's (1961) pioneering behavioral work, electrophysiological studies typically employ a paradigm in which a monkey chooses between two alternative responses that may differ with respect to the sensory information available on each trial, the prior odds, and the outcome assigned to each response alternative. These experiments yield insight into how sensory information is integrated with reward information accumulated across previous trials. Reward is typically manipulated by assigning variable amounts of juice to different color-coded response alternatives (Platt and Glimcher, 1999; Sugrue et al., 2004; Glimcher, 2008, p. X-X; Newsome et al., 2008, p. X-X; Sanfey and Dorris, 2008, p. X-X). When rewards were assigned stochastically, the monkey's choices appeared to be based on an estimate of the probability of reward gained by sampling over the last few trials (Sugrue et al., 2004). These results indicate that the brain quickly reaches a decision based on the reward history of the last few trials.

Single-cell activity in response to stochastic variations of reward has been found in ventral midbrain areas (Fiorillo et al., 2003). These dopamine neurons' phasic activity correlates with the so-called prediction error, i.e. with the difference between actual and expected reward (Morris et al., 2004; Schultz et al., 1997; see also "Section 4. Understanding Valuation – Learning Valuations", X-X) and modulating that activity affected choice (Pessiglione *et al.*, 2006). However, these same neurons also produced a tonic response that was highest in conditions of highest risk, i.e. in trials in which the probability of receiving the reward was 0.5. The behavioral relevance of this midbrain dopaminergic single-cell activity recorded in response to changes in reward probability remains controversial (Bayer and Glimcher, 2005; Morris et al., 2006; Niv et al., 2006).

Using fMRI techniques in humans, a variety of subcortical and cortical areas have been implicated in the coding of decision variables such as expected gain, probability of reward, risk and ambiguity. Most studies employ a visual representation of a gambling task and, after a delay of several seconds, subjects are instructed to choose between pairs of options by key-press. Brain activity is monitored during the delay period and correlated with various decision variables (see, e.g. Daw and Doya, 2006; Glimcher and Rustichini, 2004; Montague et al., 2006; O'Doherty, 2004; Rorie and Newsome, 2005; Trepel et al., 2005). Neural activity may be correlated with gain or loss of a potential reward, the probability of reward, their product (expected gain) or risk (the variance of gain). A number of studies suggest that reward value is encoded in the striatum and portions of prefrontal cortex (PFC) and orbitofrontal cortex (OFC) (Daw et al., 2006; Knutson et al., 2005; O'Doherty, 2004; Tanaka et al., 2005; Tom et al., 2007).

In humans, reward-prediction error signals are generally found to be localized to the striatum, although also seen in OFC and amygdala (Daw et al., 2006; O'Doherty, 2004;

Pessiglione et al., 2006; Yacubian et al., 2006). It has been difficult to disentangle probability of reward from expected gain, and most studies find responses correlated with expected gain in the striatum, OFC and medial PFC (Daw et al., 2006; Delgado et al., 2005; Hsu et al., 2005; Knutson et al., 2005; Preuschoff et al., 2006). Several of these studies see an increase in activity with increasing expected value whether or not the outcome is an expected loss or gain. On the other hand, Yacubian et al. (2006) suggest that while expected gains are encoded in the striatum, expected losses result in responses in the amygdala, perhaps also associated with negative emotion. This is supported by the finding that decisions consistent with framing effects are correlated with increased response in the amygdala. PFC response is higher in subjects that have less of a framing effect (De Martino et al., 2006) suggesting a requirement for cognitive control to suppress this cognitive bias.

An important distinction in these gambling tasks is between risk (in which the probabilities of the outcomes are known precisely) and ambiguity (when they are not). Responses correlated with risk have been found in anterior insula, OFC and striatum (Preuschoff et al., 2006), in the ventral striatum and anterior insula (Kuhnen and Knutson, 2005), as well as dorsolateral PFC and posterior parietal cortex (Huettel et al., 2005).

Responses correlated with the ambiguity of a decision have been found in the posterior part of the inferior frontal sulcus (Huettel et al., 2006), OFC, amygdala and dorsomedial PFC, along with a negative correlation with responses in the striatum (Hsu et al., 2005). Subjects with a preference for ambiguity over risk show stronger responses in lateral PFC, while subjects with a preference for risk over ambiguity show stronger responses in posterior parietal cortex (Huettel et al., 2006). A reward received immediately is generally valued more than one that will be delayed, a phenomenon known as temporal discounting. If a reward will be received

immediately, a variety of brain areas respond including striatum and OFC, but the inclusion of a potentially delayed reward recruits other areas including portions of the PFC (Glimcher et al., 2007; McClure et al., 2004; Tanaka et al., 2005) suggesting the need for cognitive control for choices involving delayed gratification.

Little is known about the neural coding of errors in pure motor tasks. Comparing errors in movement completion (induced by target displacements) with kinematic errors (induced by novel visual feedback), and dynamic errors (induced by the application of force fields) showed increased cerebellar activity both for kinematic or dynamic errors (Diedrichsen et al., 2005). Target errors, but not execution errors, activated the posterior superior parietal lobule and the striatum. In contrast, execution errors produced strong adaptive responses that specifically activated anterior aspects of the parietal cortex and the dorsal premotor cortex. Overall, structures involved in the correction of errors attributable to misestimation of dynamics were generally a subset of the neural areas involved in correction of movement errors attributable to misestimation of kinematics.

Conclusion

We have presented results from a variety of different approaches directed at understanding the processes underlying decision making in motor tasks. The results presented here indicate that movement planning shares the same formal structure as economic decision making. Subjects in movement tasks are generally found to be very good at choosing motor strategies that come close to maximizing expected gain. In contrast, subjects in economic decision making typically fail to maximize expected gain. Moreover, the sources of uncertainty in motor tasks are endogenous: they reflect the organism's own uncertainty in planning movement while, in contrast, uncertainty in economic tasks is typically imposed by the

experimenter. Thus, probabilistic information from cognition, perception, and movement has different origins.

In economic decision tasks, feedback about outcomes typically reduces biases and misperceptions in the representation of probability estimates, moving behavior closer to strategies maximizing expected gain. We emphasize that in movement planning under risk subjects' performance is initially close to optimal performance, maximizing expected gain, and does not appear to change with feedback. Movement planning is well described by simple models that maximize expected gain while there is no single model of economic decision making that captures all of the complexity of human behavior. Careful study of the neural circuitry underlying decision making in the form of movement could lead to a better understanding of how the brain gathers information to make decisions and transforms them into movement.

References

- Alexander, R. M. (1997). A minimum energy cost hypothesis for human arm trajectories. *Biol. Cybern.* **76**, 97-105.
- Allais, M. (1953). Le comportement de l'homme rationnel devant la risque: critique des postulats et axiomes de l'école Américaine. *Econometrica* **21**, 503-546.
- Attneave, F. (1953). Psychological probability as a function of experienced frequency. *J. Exp. Psychol.* **46**, 81-86.
- Baddeley, R.J., Ingram, H. A., and Miall, R. C. (2003). System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *J. Neurosci.* **7**, 3066-3075.
- Battaglia, P. W., and Schrater, P. R. (2007). Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *J. Neurosci.* **27**, 6984-6994.
- Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129-141.
- Bays, P. M., and Wolpert, D. M. (2006). Computational principles of sensorimotor control that minimize uncertainty and variability. *J. Physiol.* **578**, 387-396.
- Bell, D. E., Raiffa, H., and Tversky, A. (Eds.) (1988). *Decision making: descriptive, normative and prescriptive interactions*. Cambridge, UK: Cambridge University Press.
- Bernoulli, D. (1738/1954). Exposition of a new theory on the measurement of risk, *Comentarii Academiae Scientiarum Imperialis Petropolitanae*. Translation published in *Econometrica*, **22**, 23-36.

- Bossaerts, P. (2008). Valuation and learning as a precondition of choice: The case of decision making under uncertainty. *In: P. W. Glimcher, C. F. Camerer, E. Fehr, and R. A. Poldrack (Eds.), Neuroeconomics: Decision Making and the Brain*. London, UK: Elsevier, p. X-X.
- Burdet, E., Osu, R., Franklin, D. W., Milner, T. E., and Kawato, M. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* **414**, 446-449.
- Camerer, C. F. (1992). The rationality of prices and volume in experimental market. *Organ. Behav. Hum. Decis. Process.* **51**, 237-272.
- Cuijpers, R. H., Smeets, J. B. J., and Brenner, E. (2004). On the relation between object shape and grasping kinematics. *J. Neurophysiol.* **91**, 2598-2606.
- Daw, N.D., and Doya, K. (2006). The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* **16**, 199-204.
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876-879.
- Dayan, P., and Balleine, B.W. (2002) Reward, motivation and reinforcement learning. *Neuron* **36**, 285-298.
- Dean, M., Wu, S.-W., and Maloney, L. T. (2007). Trading off speed and accuracy in rapid, goal-directed movements, *J. Vis.* **7**, 1-12.
- Delgado, M. R., Miller, M. M., Inati, S., and Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *NeuroImage* **24**, 862-873.
- De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* **313**, 684-687.

- Diedrichsen J., Hashambhoy Y., Rane T., and Shadmehr R. (2005). Neural correlates of reach errors. *J. Neurosci.* **25**, 9919-9931.
- Dornay, M., Uno, Y., Kawato, M., and Suzuki, R. (1996). Minimum muscle-tension change trajectories predicted by using a 17-muscle model of the monkey's arm, *J. Mot. Behav.* **2**, 83-100.
- Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898-1902.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* **47**, 381-391.
- Flash, T., and Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model, *J. Neurosci.* **5**, 1688-1703.
- Franklin, D. W., Liaw, G., Milner, T. E., Osu, R., Burdet, E., and Kawato, M. (2007). Endpoint stiffness of the arm is directionally tuned to instability in the environment. *J. Neurosci.* **27**, 7705-7716.
- Gepshtein, S., Seydell, A., and Trommershäuser, J. (2007). Optimality of human movement under natural variations of visual-motor uncertainty. *J. Vis.* **7**, 1-18.
- Glimcher, P. W., and Rustichini, A. (2004). Neuroeconomics: The consilience of brain and decision. *Science* **306**, 447-452.
- Glimcher, P. W., Kable, J., and Louie, K. (2007). Neuroeconomic studies of impulsivity: Now or just as soon as possible? *American Economic Review*, in press.
- Glimcher, P. (2008). From value to choice. In: P. W. Glimcher, C. F. Camerer, E. Fehr, and R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*. London, UK: Elsevier, p. X-X.

- Hamilton, A. F. C., and Wolpert, D. M. (2002). Controlling the statistics of action: obstacle avoidance, *J. Neurophysiol.* **87**, 2434-2440.
- Harris, C. M., and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature* **394**, 780-784.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4**, 267-272.
- Hertwig, R., Barron, G., Weber, E. U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* **15**, 534-349.
- Holt, C., and Laury, S. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* **92**, 1644-1655.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* **310**, 1680-1683.
- Hudson, T. E., Maloney, L. T., and Landy, M. S. (2007). Optimal movement timing with temporally asymmetric penalties. Under review.
- Huettel, S.A., Song, A. W., and McCarthy, G. (2005). Decisions under uncertainty: Probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* **25**, 3304-3311.
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., and Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* **49**, 765-775.
- Kahneman, D., Slovic, P., and Tversky, A. (Eds.) (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kahneman, D., and Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica* **47**, 263-292.

- Kahneman, D., and Tversky, A. (Eds.) (2000). *Choices, values and frames*. New York: Cambridge University Press.
- Kaminsky, T., and Gentile, A. M. (1986). Joint control strategies and hand trajectories in multijoint pointing movements, *J. Mot. Behav.* **18**, 261-278.
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005). Distributed neural representation of expected value. *J. Neurosci.* **25**, 4806-4812.
- Körding, K. P., Ku, S. P., and Wolpert, D. M. (2004). Bayesian integration in force estimation. *J. Neurophysiol.* **92**, 3161-3165.
- Körding, K. P., and Wolpert, D. M. (2004a). Bayesian integration in sensorimotor learning. *Nature* **427**, 244-247.
- Körding, K. P., and Wolpert, D. M. (2004b). The loss function of sensorimotor learning. *Proc. Natl. Acad. Sci. U S A* **101**, 9839-9842.
- Kuhnen, C. M., and Knutson, B. (2005). The neural basis of financial risk taking. *Neuron* **47**, 763-770.
- Landy, M. S., Goutcher, R., Trommershäuser, J., and Mamassian, P. (2007). Visual estimation under risk. *J. Vis.* 7:4, 1-15.
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978). Judged frequency of lethal events. *J. Exp. Psychol. [Hum. Learn.]* **4**, 551-578.
- Maloney, L. T., Trommershäuser, J., and Landy, M. S. (2007). Questions without Words: A Comparison between Decision Making under Risk and Movement Planning under Risk. In: *Integrated Models of Cognitive Systems* (W. Gray, Ed.), pp. 297-315. New York: Oxford University Press.

- Ma-Wyatt, A., Stritzke, M., and Trommershäuser, J. (2007). Eye-hand coordination can be altered for rapid pointing under risk. *J. Vis*, under review.
- McClure, S. M., Laibson, D. I., Lowenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* **306**, 503-507.
- Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., and Smith, J. E. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychol. Rev.* **95**, 340-370.
- Montague, P. R., King-Casas, B., and Cohen, J. D. (2006). Imaging valuation models in human choice. *Annu. Rev. Neurosci.* **29**, 417-448.
- Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* **43**, 133-143.
- Newsome, W., Sugrue, L., and Corrado, G. (2008). The trouble with choice: Finding decision variables in the brain *In*: P. W. Glimcher, C. F. Camerer, E. Fehr, and R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*. London, UK: Elsevier, p. X-X.
- Niv, Y., Daw, N. D., and Dayan, P. (2006). Choice value. *Nature Neuroscience* **9**, 987-988.
- O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* **14**, 769-776.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042-1045.
- Platt, M. L. and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233.

- Preuschoff, K., Bossaerts, P., and Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**, 381-390.
- Redelmeier, D. A., and Tversky, A. (1992). On the framing of multiple prospects. *Psychol. Sci.* **3**, 191-193.
- Rorie, A. E., and Newsome, W. T. (2005). A general mechanism for decision-making in the human brain? *Trends Cogn. Sci.* **9**, 41-43.
- Sabes, P. N., and Jordan, M. I. (1997). Obstacle avoidance and a perturbation sensitivity model for motor planning. *J. Neurosci.* **17**, 7119-7128.
- Sanfey, A., and Dorris, M. (2008). Games in humans and non-human primates: From scanners to single units. In: P. W. Glimcher, C. F. Camerer, E. Fehr, and R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain*. London, UK: Elsevier, p. X-X.
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* **275**, 1593-1599.
- Sedlmeier, P., Hertwig, R., and Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *J. Exp. Psychol. Learn. Mem. Cogn.* **24**, 754-770.
- Soechting, J. F., Buneo, C. A., Herrmann, U., and Flanders, M. (1995). Moving effortlessly in three dimensions: Does Donders' Law apply to arm movement? *J. Neurosci.* **15**, 6271– 6280.
- Soechting, J. F., and Lacquaniti, F. (1981). Invariant characteristics of a pointing movement in man. *J. Neurosci.* **1**, 710–720.
- Stritzke, M., and Trommershäuser, J. (2007). Rapid visual localization during manual pointing under risk. *Vision Res.* **47**, 2000-2009.

- Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782-1787.
- Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**, 363-375.
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*, Cambridge, MA: MIT Press.
- Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2005). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* **7**, 887-893.
- Tassinari, H., Hudson, T. E., and Landy, M. S. (2006). Combining priors and noisy visual cues in a rapid pointing task. *J. Neurosci.* **26**, 10154-10163.
- Thaler, R., and Johnson, E.J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Manage. Sci.* **36**, 643-660.
- Todorov, E. (2004). Optimality principles in sensorimotor control. *Nat. Neurosci.* **9**, 907-915.
- Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* **315**, 515-518.
- Trepel, C., Fox, C. R., and Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Brain Res. Cogn. Brain Res.* **23**, 34-50.
- Trommershäuser, J., Gepshtein, S. G., Maloney, L. T., Landy, M. S., and Banks, M. S. (2005). Optimal compensation for changes in task-relevant movement variability. *J. Neurosci.* **25**, 7169-7178.
- Trommershäuser, J., Landy, M. S., and Maloney, L. T. (2006a) Humans rapidly estimate expected gain in movement planning. *Psychol. Sci.* **17**, 981-988.

- Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2003a). Statistical decision theory and trade-offs in the control of motor response. *Spat. Vis.* **16**, 255-275.
- Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2003b). Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. A* **20**, 1419-1433.
- Trommershäuser, J., Mattis, J., Maloney, L. T., and Landy, M. S. (2006b). Limits to human movement planning with delayed and unpredictable onset of needed information. *Exp. Brain Res.* **175**, 276-284.
- Tversky, A., and Kahneman, D. (1992). Advances in Prospect Theory: Cumulative representation of uncertainty. *J. Risk Uncertain.* **5**, 297-323.
- Uno, Y., Kawato, M., and Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model. *Biol. Cybern.* **61**, 89–101.
- von Neumann, J., and Morgenstern, O. (1944). *The Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.
- Wakker, P. P., Thaler, R. H., and Tversky, A. (1997). Probabilistic insurance. *J. Risk Uncertain.* **15**, 7-28.
- Wu, S.-W., Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2006). Limits to human movement planning in tasks with asymmetric value landscapes. *J. Vis.* **5**, 53-63.
- Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., and Büchel, C. (2006). Dissociable systems for expected gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* **26**, 9530-9537.

Figure Captions

Fig. 1. Example of a model of movement planning as optimization of a biomechanical constraint.

(A) Subjects moved a manipulandum between various target positions (labeled T1, ..., T6). (B) Hand path toward one example target. (C) Hand velocity. (D-E) Acceleration in the x (panel D) and y (panel E) directions. Dashed curves: mean measured hand path. Solid curves: hand paths predicted by a model that minimized integrated jerk (rate of change of acceleration). The measured hand path shows a typical smoothly curved trajectory with a bell-shaped velocity profile. The minimum-jerk model (dashed curves) does a reasonable job of predicting movement kinematics. Figure reprinted from Flash and Hogan (1985).

Fig. 2. Predictions of a model of a 10 deg, 50 ms eye movement with signal-dependent noise.

Dashed curves: the result of “bang-bang” control in which nearly maximum acceleration is applied throughout the movement until just prior to movement completion, at which point the eye is quickly brought to a stop. Solid curves: movement achieving minimum positional variance at the end of the movement. The minimum-variance model predicts a bell-shaped velocity profile and returns in substantially lower end-point variance. (A) Control signal. (B) Position. (C) Velocity, (D) Positional variance. Figure reprinted from Harris and Wolpert (1998).

Fig. 3. A movement task equivalent to decision making under risk. Subjects were required to touch a computer screen within a short period of time (e.g. 700 ms). Subjects won 100 points by hitting inside the green circle and lost 500 points by hitting inside the red circle. Subjects did not win or lose points by hitting the background, as long as the movement was completed before the time limit, but a large penalty (700 points) was imposed for movements that were too slow.

Because movements were speeded, they were also variable: the finger did not always land where the subject intended. As a result, it was in the best interest of the subject to aim to the right of the center of the green circle, compromising between the loss of points due to occasionally missing the green target circle and the loss of points due to occasionally landing in the red penalty circle.

Fig. 4. Equivalence between movement under risk and decision making under risk. (A) There are four regions in which the end point can land in the task outlined in Fig. 3: reward-only (region R_1 with expected gain +100), reward-penalty overlap (region R_2 with expected gain -400), penalty-only (region R_3 with expected gain -500) and background (region R_4 with expected gain 0). (B) Sample end points for a subject aiming at the center of the green target (aim point indicated by the white diamond). This subject had a motor uncertainty of 5.6 mm (standard deviation); target and penalty circles have radii of 9 mm. The combination of aim point and motor uncertainty specifies the probability of landing in each region. This movement strategy yields an expected gain of -111.39 points/trial due to the high frequency of touching inside the red penalty circle. (C) Sample end points for the same subject aiming to the right of the target center to avoid accidental hits inside the penalty circle. The expected gain = 20.68 points/trial corresponds to the optimal strategy maximizing expected gain. (D) Expected gain for this subject as a function of mean movement end point. The maximum of this function corresponds to the aim point illustrated in (C).

Fig. 5 Movement strategies during rapid movement under risk. Measured mean movement end point as a function of optimal end point maximizing expected gain (data reprinted from Trommershäuser et al., 2003b) in the task illustrated in Fig. 3. Five subjects participated. There

were six different conditions corresponding to green target circles located to the left (open symbols) or right (filled symbols) of the red penalty circle at each of three different target-penalty distances. The data points fall close to the diagonal identity line, indicating that subjects chose aim points close to those maximizing expected gain based on their individual movement variances.

Fig. 6. Movement timing under risk. (A) In this task, subjects reached toward a target and were rewarded if they hit the target arriving within a fixed range of movement times (as indicated by the green portions of these timer bars). In four different conditions, either early or late movements arriving within a second range of movement times were penalized (the red portions of the timer bars). (B) Calculation of expected gain as a function of mean movement duration. Upper panel: distribution of movement times for several mean durations. Note that variance increases with mean duration. Middle panel: expected gain as a function of actual movement duration for one condition. Bottom panel: expected gain varies with mean movement duration; the black circle denotes the mean duration resulting in maximum expected gain. (C) Predictions for one subject; the four panels correspond to the four conditions in (A). Black circles indicate the MEG mean temporal aim point. Diamonds indicate observed mean durations. Across the four conditions and five subjects, mean movement durations were close to those that would result in maximum expected gain (Hudson et al., 2007).

Fig. 7. Calculation of expected gain in the experiment of Dean et al. (2007). Subjects reached for a target on a display screen and were rewarded if they hit the target. The value of a hit on the target decreased linearly with movement duration (dashed line). Increased movement duration

results in decreased movement end-point variance, leading to an increased probability of a hit on the target (solid curve). The product of the value and probability of a hit equals expected value (dotted curve). The duration corresponding to maximum expected gain is indicated by the circle.

Fig. 8. Task and calculation of the optimal movement strategy in the experiment of Battaglia and Schrater (2007). (A) Targets were indicated by dots drawn from a Gaussian distribution centered on the unseen target. Dots were drawn one by one at a constant rate and no more dots were drawn once the subject began to reach, so that subjects had more information concerning target location the longer they waited to move. “S” indicates the starting point of the reach and the unseen target was located on the circular arc. The black bar indicated the amount of time that had elapsed throughout each trial. (B) Hits on the target were rewarded only if they occurred within 1200 ms from the display of the first dot. The x -axis indicates the time of movement onset. Larger movement-onset times leave less time available for the movement, resulting in greater standard deviation of end point location due to motor variability (dotted curve) but smaller standard deviation due to error in visual estimation of target location (dashed curve). The overall standard deviation (relative to the target center) determines the probability of reward. The minimal standard deviation (corresponding to the strategy maximizing expected gain) is indicated. This varied with the standard deviation of the probability from which the dots were sampled (“dot scatter level”) as shown in the three panels.

Fig. 9. Consistency of pointing strategy across trials. (A) Trial-by-trial deviation of movement end point from mean movement end point as a function of trial number after introduction of rewards and penalties; the six different lines correspond to the six different spatial conditions of

target and penalty offset (data replotted from Fig. 7, Trommershäuser et al., 2003a). (B) Trend of a hypothetical learning model in which a subject changes motor strategy gradually in response to rewards and penalties incurred. The subject initially aims at the center of the green circle. Before the subject's first trial in the decision-making phase of the experiment, the subject is instructed that red circles carry penalties and green circles carry rewards. Subjects may approach the optimal aim point maximizing expected gain by slowly shifting the aim point away from the center of the green circle until the winnings match the maximum expected gain. However, the data shown in (A) do not exhibit this trend and do not support such a learning model.

A

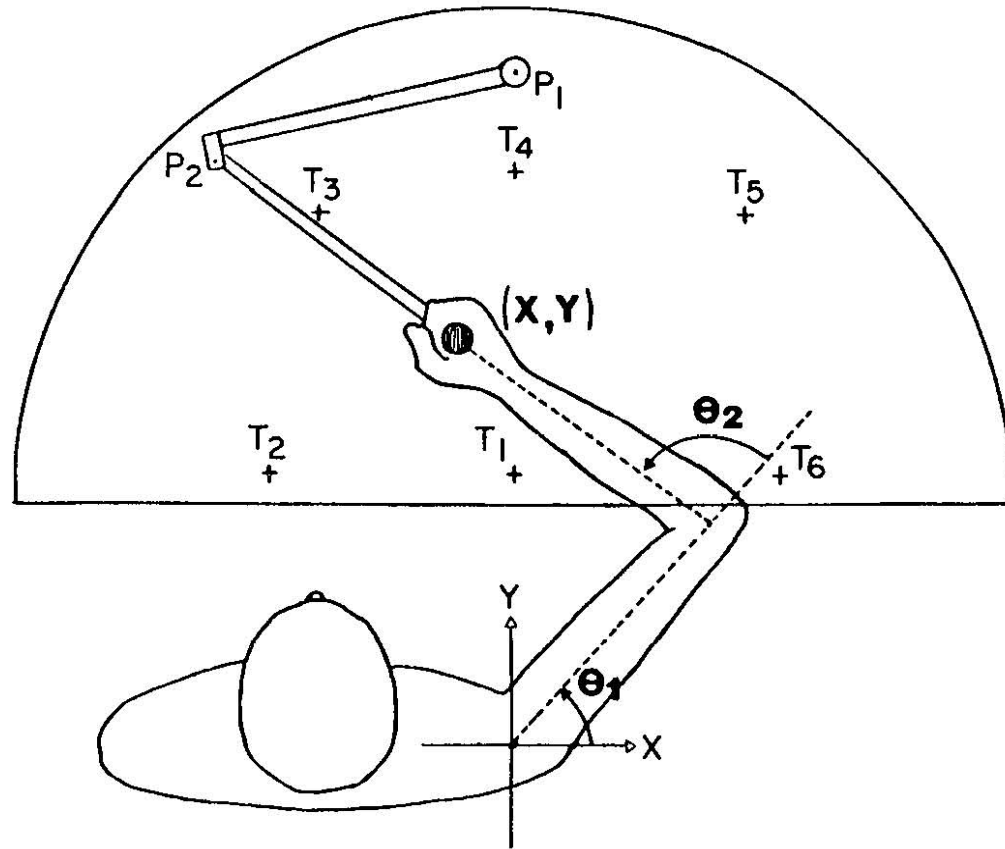


Figure 1 A

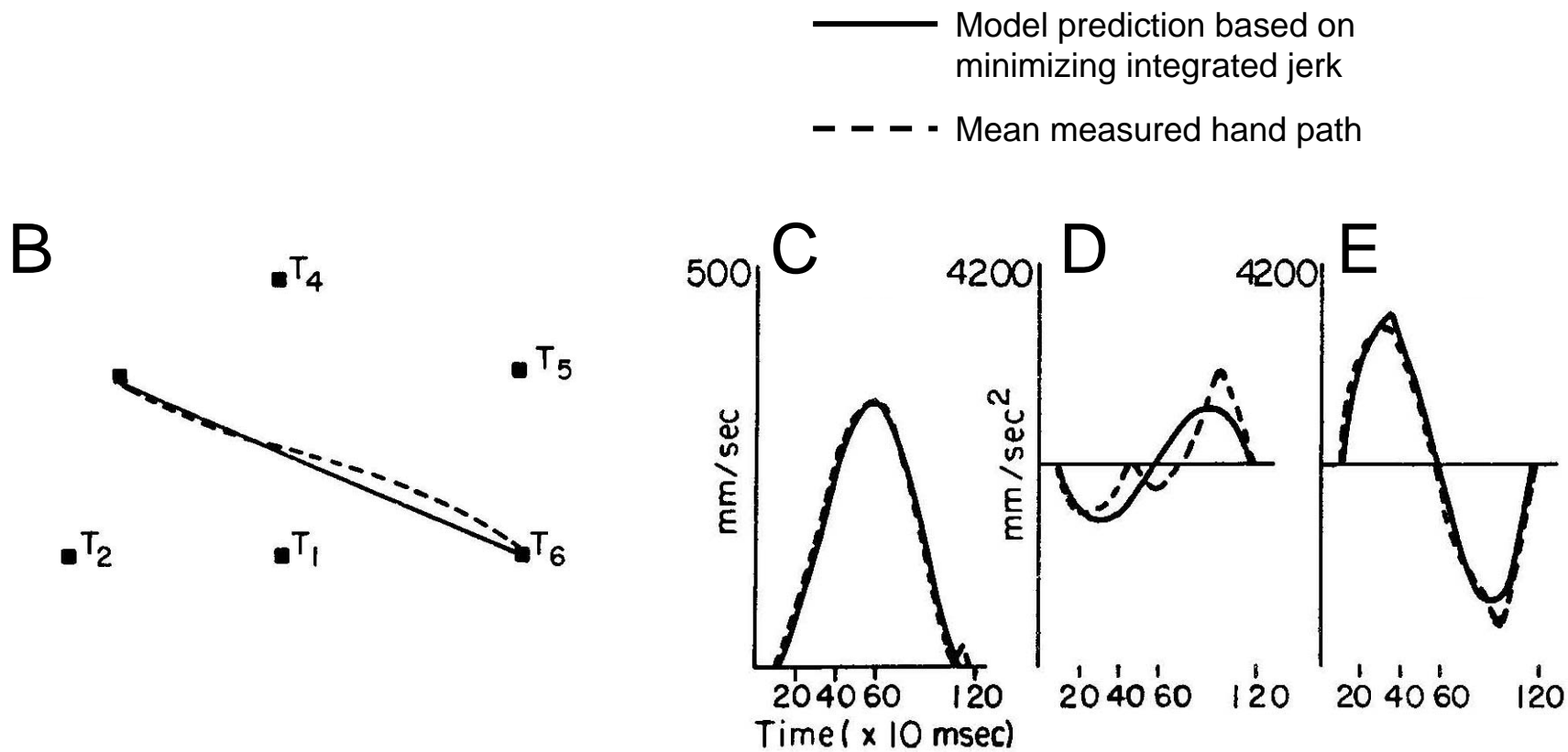


Figure 1 B-E

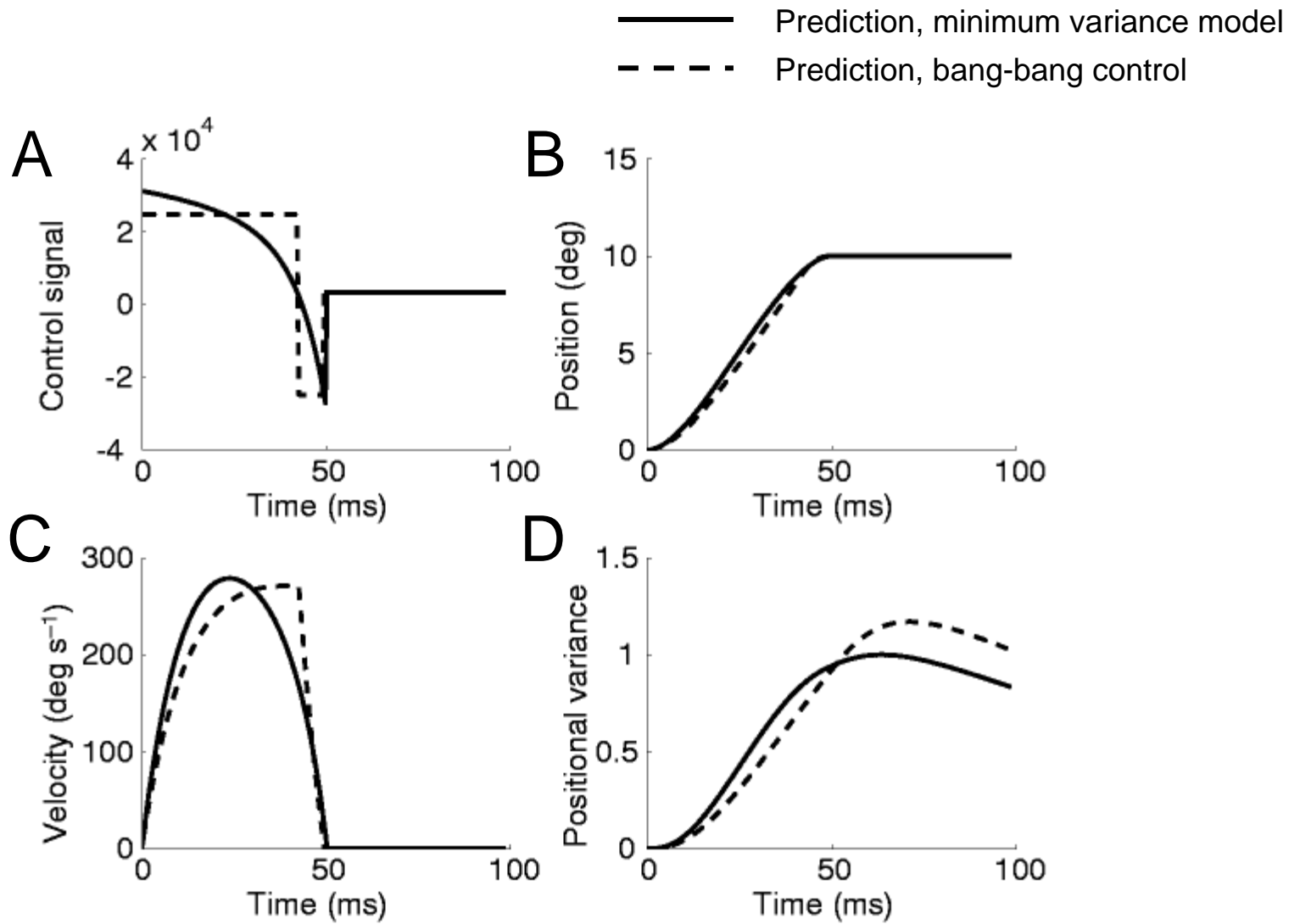


Figure 2 A-D

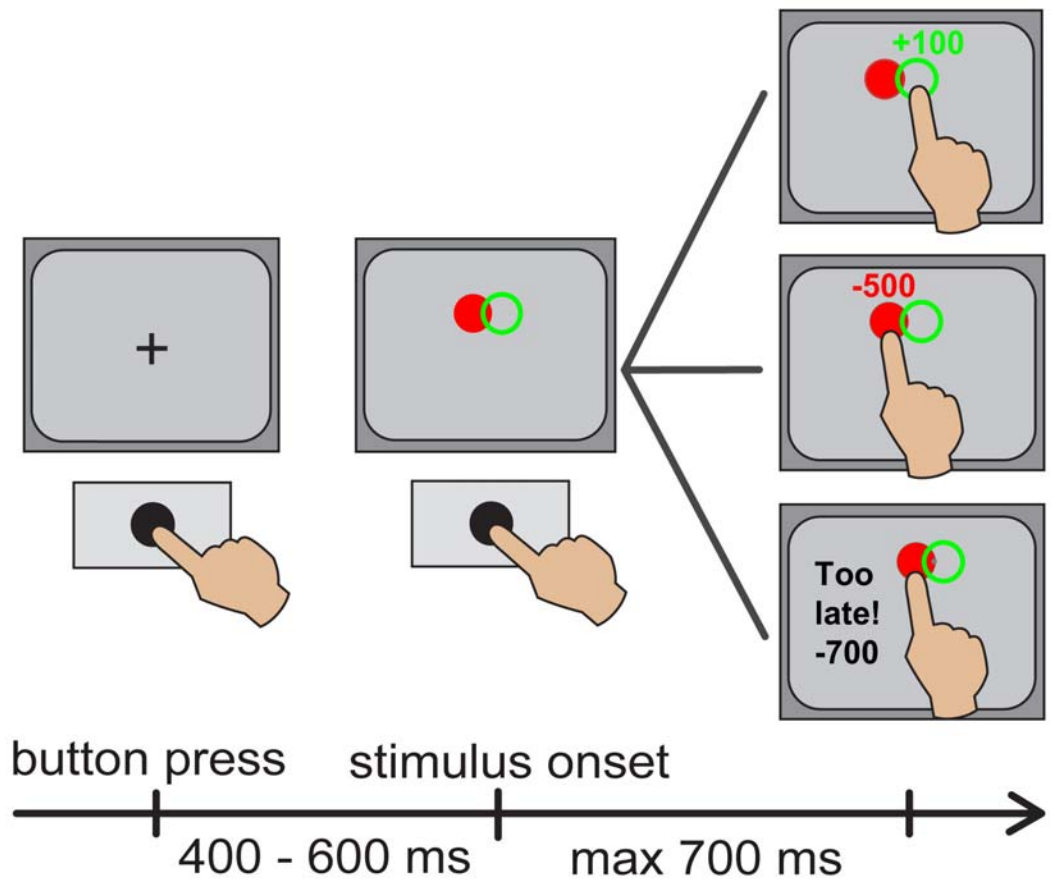


Figure 3

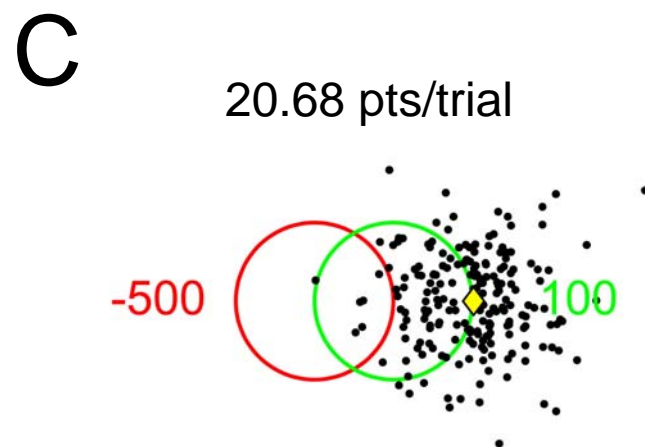
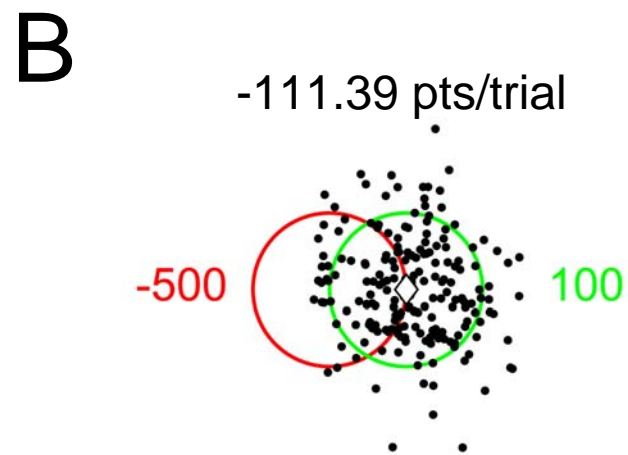
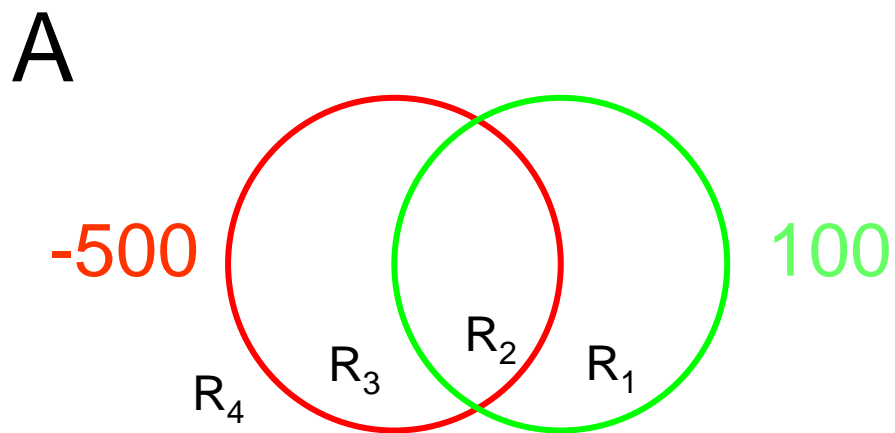


Figure 4 A-C

D

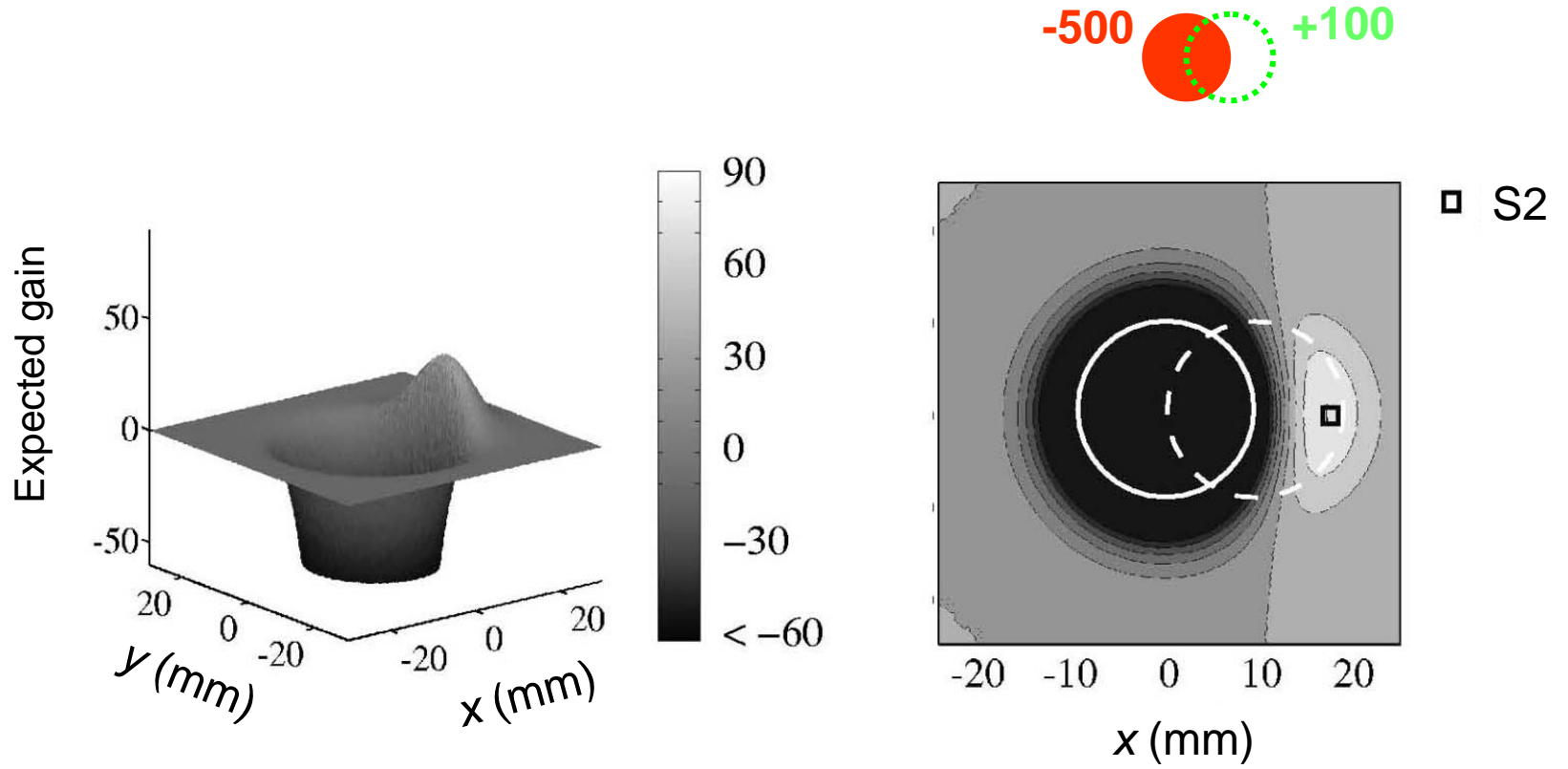


Figure 4 D

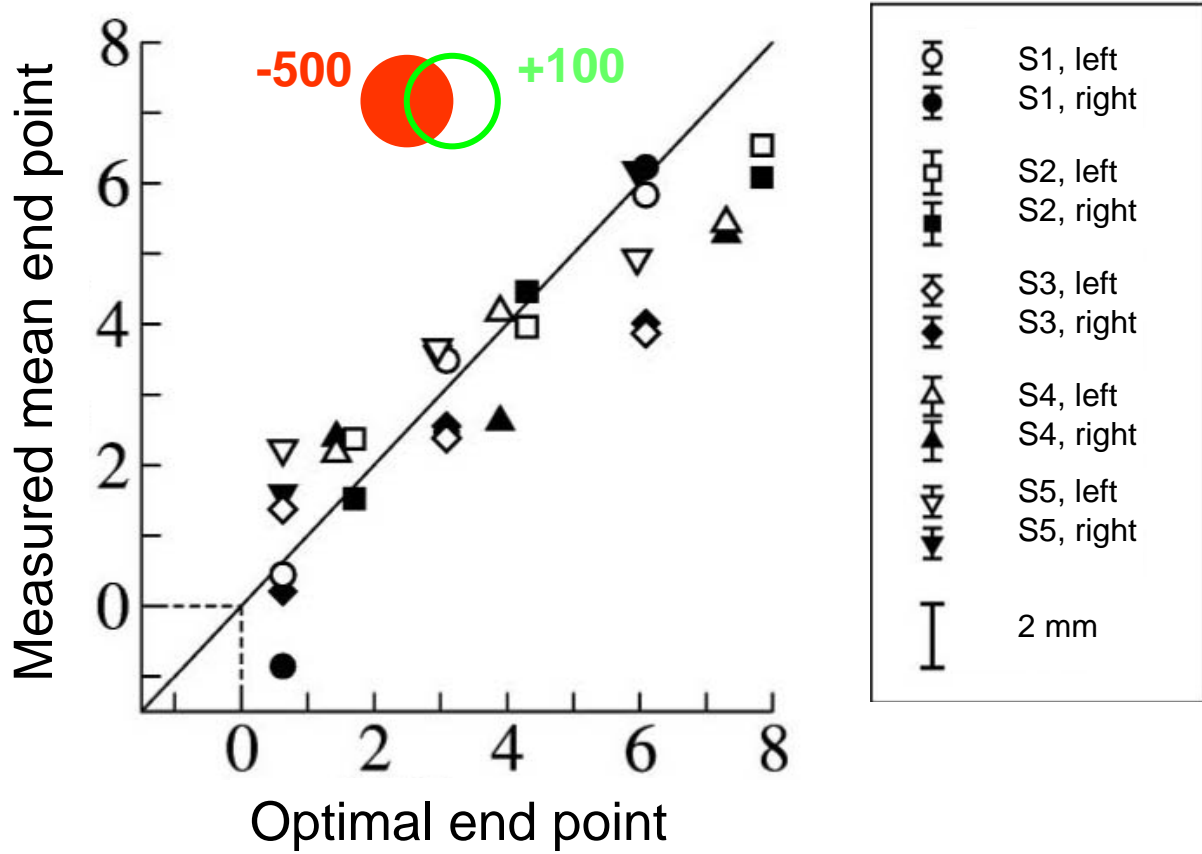


Figure 5

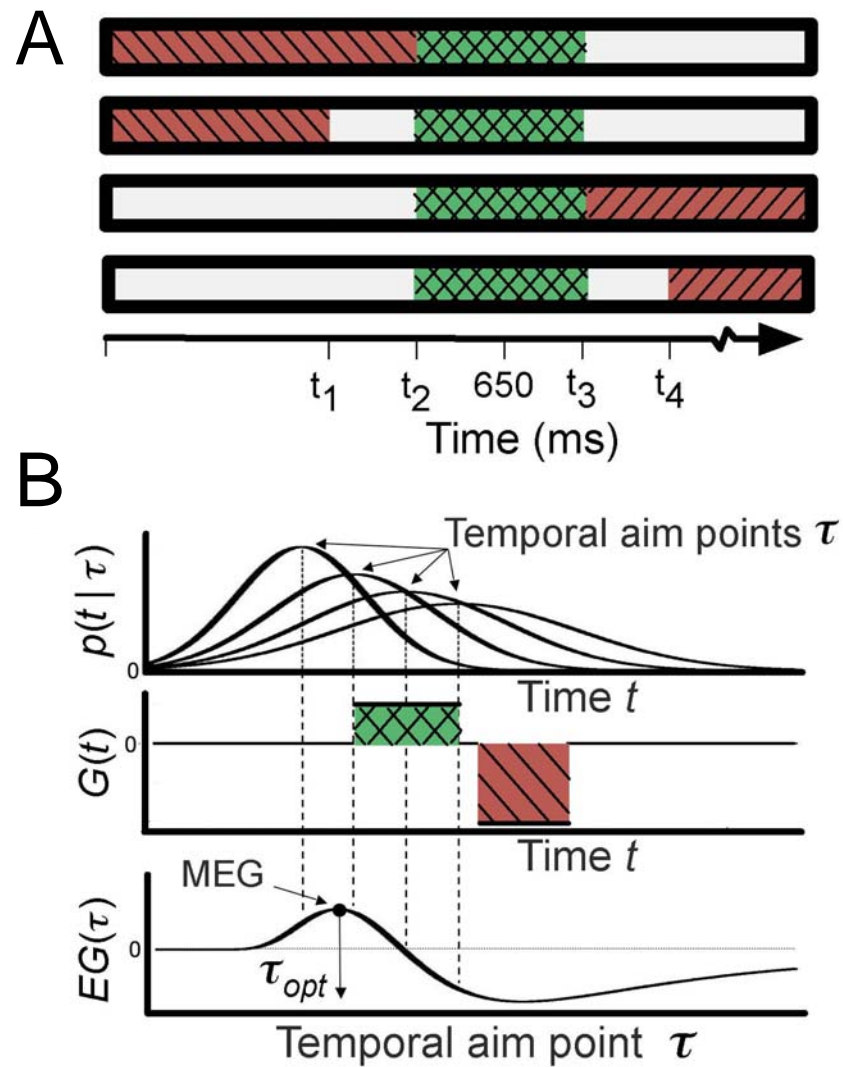
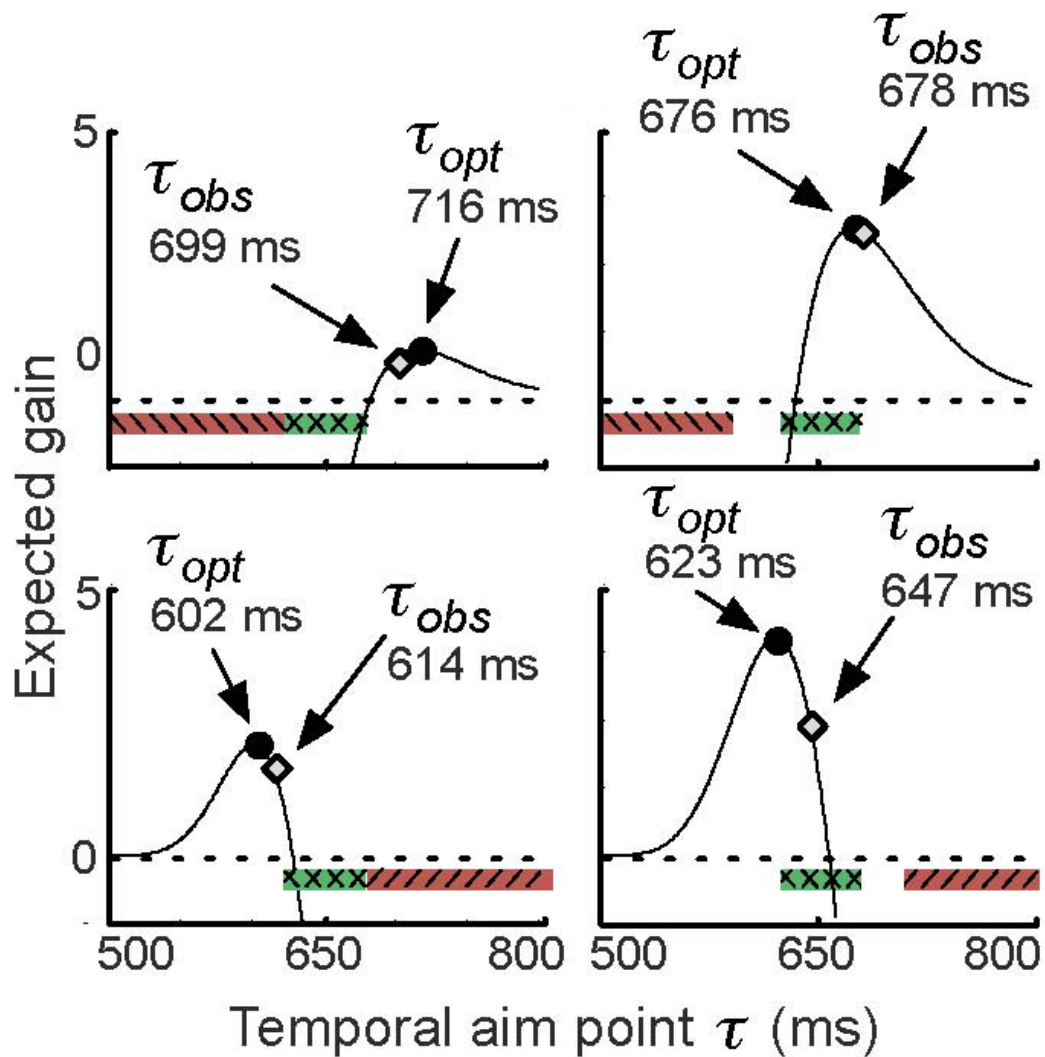


Figure 6 A-B

C**Figure 6 C**

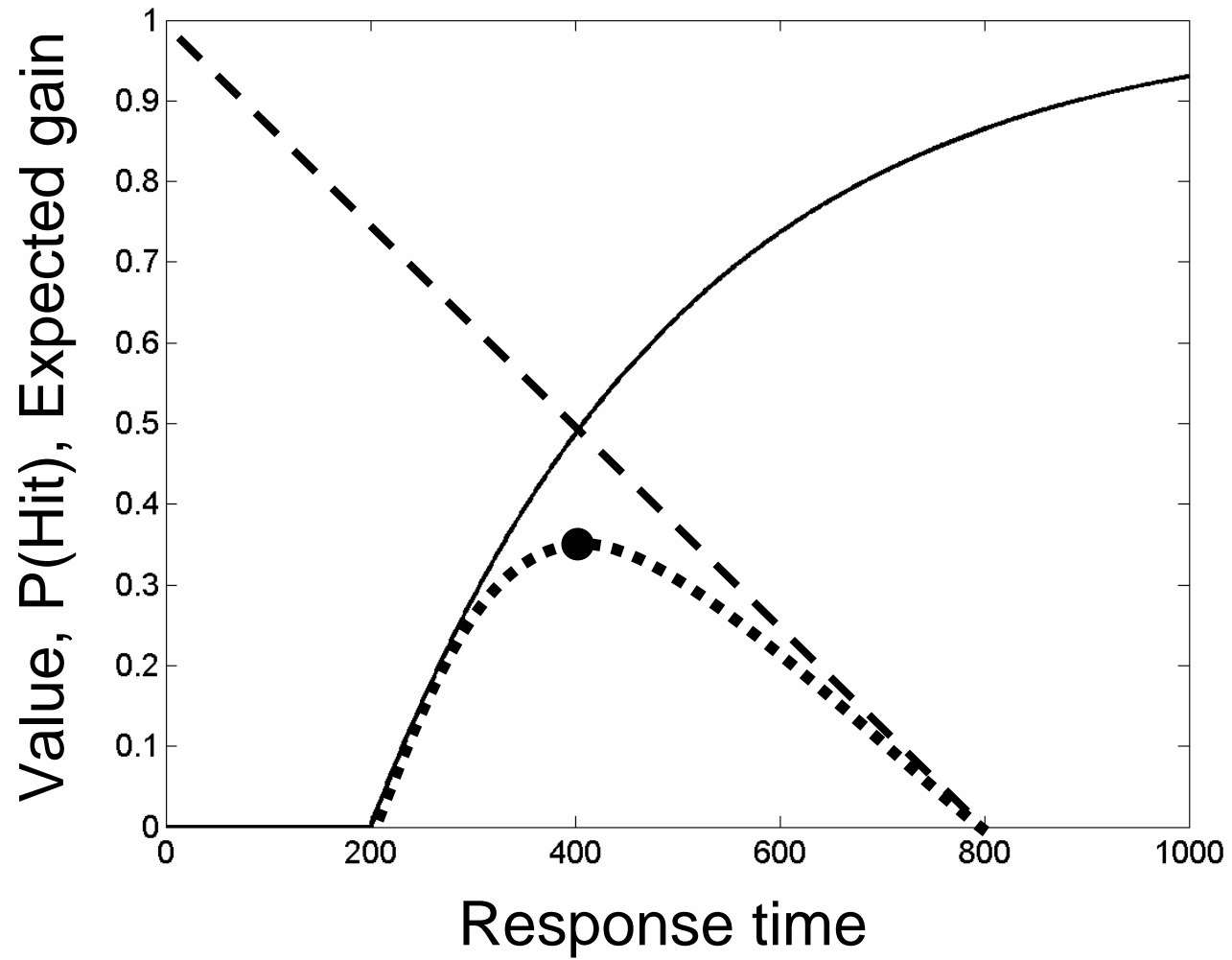


Figure 7

A

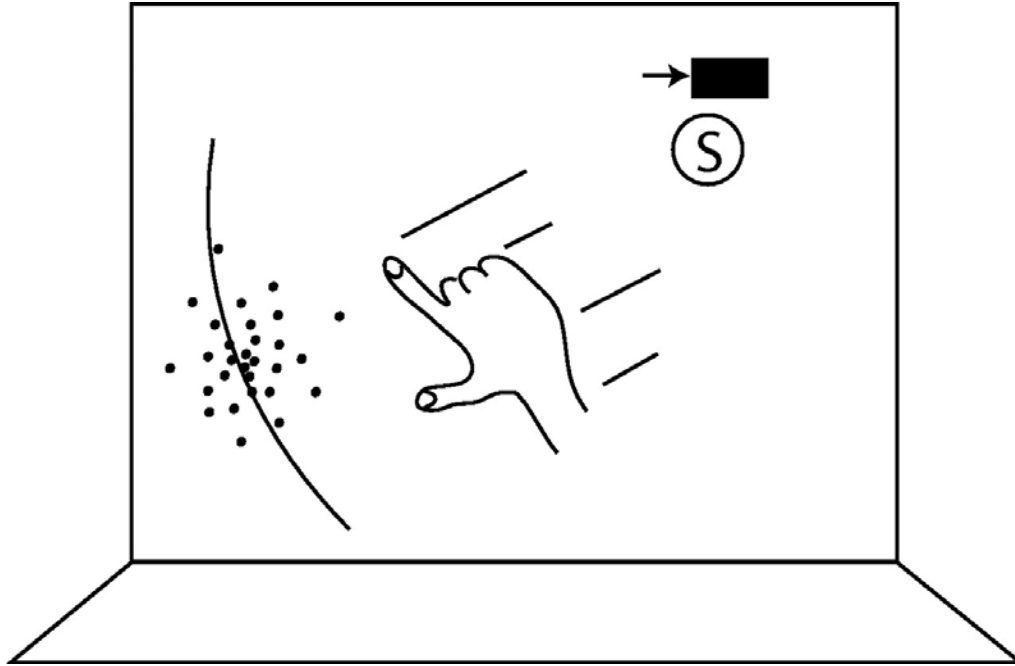


Figure 8 A

B

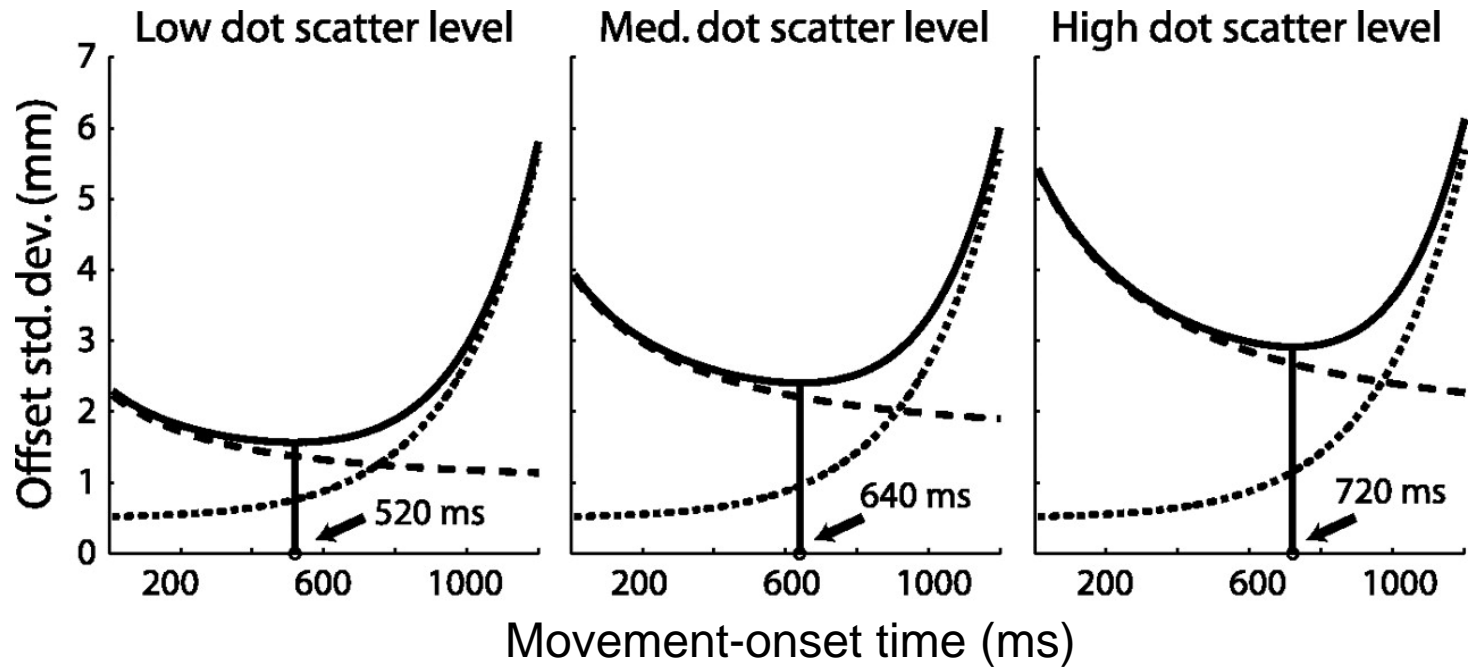


Figure 8 B

A

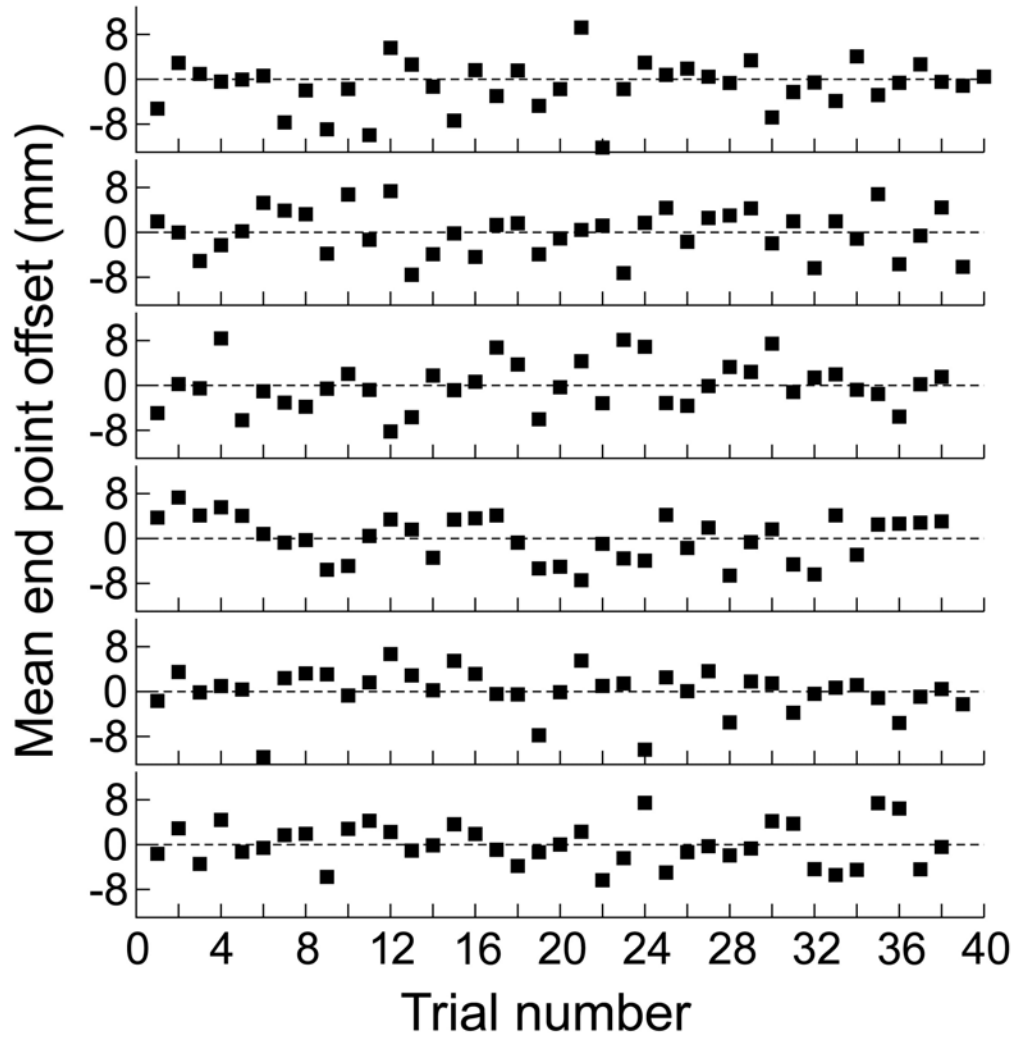


Figure 9 A

B

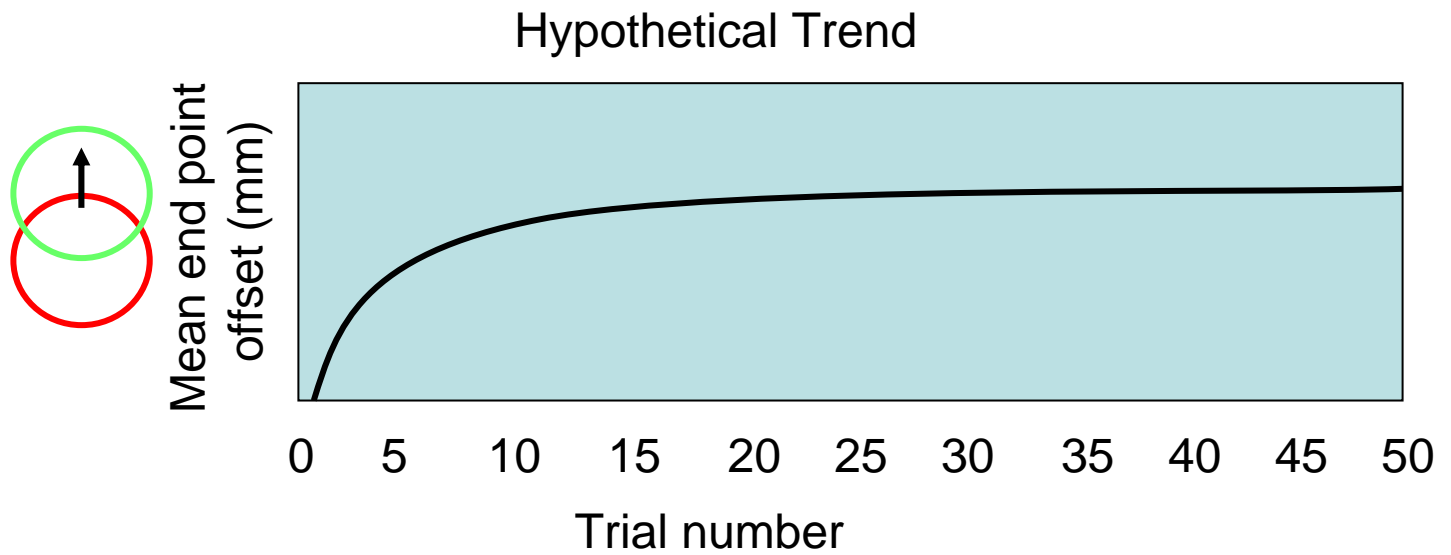


Figure 9 B