Statistical decision theory and the selection of rapid, goal-directed movements

Julia Trommershäuser, Laurence T. Maloney, and Michael S. Landy

Department of Psychology and Center for Neural Science, New York University, New York, New York 10003

Received October 1, 2002; revised manuscript received January 30, 2003; accepted February 3, 2003

We present two experiments that test the range of applicability of a movement planning model (MEGaMove) based on statistical decision theory. Subjects attempted to earn money by rapidly touching a green target region on a computer screen while avoiding nearby red penalty regions. In two experiments we varied the magnitudes of penalties, the degree of overlap of target and penalty regions, and the number of penalty regions. Overall, subjects acted so as to maximize gain in a wide variety of stimulus configurations, in good agreement with predictions of the model.
OCIS codes: 330.4060, 330.7310.

1. INTRODUCTION

Motor responses have consequences. In the first semifinal of the 2002 World Cup, Germany met host South Korea, and the match was decided in the 75th minute by the only goal of the match. Oliver Neuville passed the ball to Germany's play maker Michael Ballack, who scored in his second attempt. Aware of Korea's defenders closing in behind him, Ballack was forced to react quickly, and he first tried to score immediately by shooting straight at the goal keeper, South Korea's Lee Woon-Jae. Lee blocked the shot but Ballack regained control of the ball. He now had a second chance, a brief window in time when he could score the decisive goal by passing Lee on the right.

Imagine you are Ballack in that instant of time, illustrated in Fig. 1(a). Your only option is to try to score by aiming somewhere to the right of Lee Woon-Jac. The further you aim to your right, away from him, the lower the probability that Lee will succeed in blocking your shot. But the further you aim to the right, the greater the chances that your shot will miss the goal altogether [Fig. 1(b)]. In making your decision, you need to balance the probability that the goal keeper will block the shot against the probability of missing the goal to the right. You also need to keep in mind that you are in the World Cup semifinals and a goal will almost certainly assure victory. If you delay the decision significantly, you will lose your chance to score. You have just a fraction of a second. Where should you aim?

In this paper we examine human performance in experimental tasks analogous to the example just presented. Following a signal, the subject is required to rapidly touch a green circle on a computer screen with his or her fingertip. If the subject hits within the circle ("the goal"), she or he wins 100 points. If the subject instead hits a red, circular penalty region near the green circle ("the goal keeper"), she or he loses 500 points. Hitting anywhere else on the screen incurs no penalty but generates no reward. There is a large penalty for failure to respond within a brief time after the signal ("the defenders will steal the ball"). Owing to the short time limit and small target region, the subject cannot be certain that a movement aimed at the center of the reward region will not, instead, end up outside the reward region, possibly in the penalty region. To summarize, the subject, in each trial, must select among possible actions and must do so very rapidly. There are explicit monetary penalties associated with the outcome of the action selected, but the outcome is never completely certain. In selecting an action, the subject must strike a balance among competing risks.

In previous work[1] we developed a model of ideal performance for "action under risk" based on statistical decision theory.[2–5] The model, called MEGaMove, *M*aximizes *E*xpected *Ga*in in its selection of *Move*ment strategies. It predicts specific shifts in a subject's mean movement end point in response to changes in the reward and penalty structure of the environment and with changes in the subject's motor variability. In an experiment, subjects did indeed alter movement end points in response to changes in the amount of penalty and location of a single penalty region. In this paper we further explore the range of validity of our model in two experiments.

MEGaMove is a model of *movement* planning. In contrast, models of *motor* planning are intended to solve a different problem. Some motor-planning models emphasize selection of a single deterministic trajectory that minimizes biomechanical costs while achieving a prespecified goal of the movement. These models differ primarily in the choice of biomechanical cost function, which may include measures of joint mobility,[6,7] muscle tension changes,[8] mean squared rate of change of acceleration,[9] mean torque change[10] and peak work.[11] In a different approach, Rosenbaum and colleagues[12] proposed that the motor system computes a goal posture based on a database of stored posture representations. The goal posture is chosen on the basis of two criteria, accuracy and efficiency, which leads to a trade-off between the accuracy of a potential goal posture and the biomechanical costs of getting there. Adding obstacles to the environment changes the range of possible movement paths.[13–17] These findings are typically explained by assuming that

1084-7529/2003/071419-15$15.00

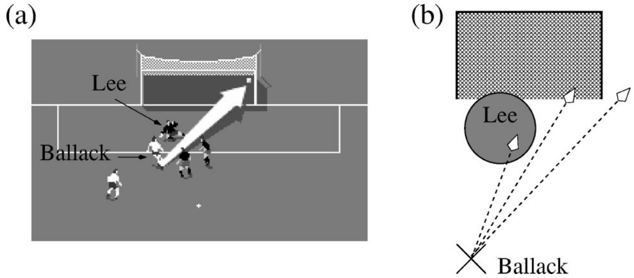

Fig. 1. Michael Ballack's goal during the 2002 World Cup. (a) Ballack, must rapidly decide where to shoot. (b) A schematic of factors affecting the decision.

trajectories are chosen to avoid spatial over-lap of the limb with the obstacle[18] or by minimizing biomechanical costs within a constrained space.[19–21]

The outcome of the models just described is a motor plan to achieve a prespecified goal, i.e., a unique sequence of motor commands selected by optimizing the trade-off between the goal of the task and biomechanical costs. These models do not take into account any error during the execution of the motor plan. Recent stochastic models of motor planning[19,22,23] rest on the assumption that the realized trajectory is affected by signal-dependent noise in the neural control signal that is driving eye and arm movements. In these models, stochastic variability is taken into account in motor planning by selecting a motor plan that minimizes the variance of the final eye or arm position or that avoids collision with an obstacle.[20,22,24] In particular, such models account for the speed–accuracy trade-off observed in rapid-pointing tasks. In these experiments subjects slow down when pointing at smaller targets to hit the target with constant reliability.[24–26]

Models of motor planning address the problem of defining and finding the optimal sequence of motor commands given a prespecified goal posture. These models do not predict the goal posture, where it is located in space, or why the movement should arrive there within a certain time window. This latter problem, movement planning, is what concerns us here.

The goal of a motor task may involve aspects other than accuracy. Experimenter-imposed (extrinsic) constraints on the task, such as monetary penalties for crossing certain regions in space, can—as long as the penalty is high enough—alter the motor plan, even if this changed motor plan results in increased motor variability. On the other hand, the consequences of following a particular motor trajectory through a "landscape" of penalties will depend critically on the motor variability associated with a given motor plan. Hence the optimal selection of a movement plan will depend on both the (monetary) reward and loss structure of the environment and the expected motor variability associated with a given motor plan. Therefore we developed the MEGaMove model which takes into account extrinsic costs associated with the task *and* the subject's own motor variability.[1]

2. A STATISTICAL DECISION-THEORY MODEL OF POINTING MOVEMENTS

Here we briefly summarize the key components of our model.[1]

1. The outcome of movement planning is a visual–motor strategy denoted S. For the simple sort of task studied in our current experiments, the choice of S will correspond to the selection of the mean movement end point within a given time limit (Section 3). However, as the word "strategy" suggests, a visual–motor strategy could also be a detailed sequence of motor commands, possibly involving intermediate goals in space and time. It could well incorporate the use of available visual or proprioceptive information to control the movement during execution.

2. When a motor strategy is executed, the result is a particular movement trajectory. In the case of our simple task, a movement trajectory $\tau(t)$ specifies, for any time t, the position of the fingertip in time and space: $\tau{:}t \rightarrow [x(t), y(t), z(t)]$. More generally, it could specify the full time course of movement in an appropriate joint-angle space.

3. When the motor system selects a particular visual–motor strategy S, it in effect imposes a probability density $P(\tau|S)$ on the space of possible movement trajectories that could occur once the motor strategy is executed. The choice of strategy influences the probability that any particular trajectory will occur, but the actual trajectory is not determined by the choice of strategy.

The probability density $P(\tau|S)$ is likely affected by the goal of the movement, the planned duration, the possibility of visual feedback during the movement,[27,28] previous training, and intrinsic uncertainty in the motor system. Computing $P(\tau|S)$ on the basis of the underlying sequence of motor commands in the presence of visual feedback and signal-dependent noise is a challenging task and has only recently been addressed within the framework of optimal feedback control theory.[23] However, note that as long as we know or are able to measure $P(\tau|S)$, the consequences of the choice of S 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 determine $P(\tau|S)$.

4. Whatever the reward/penalty structure of the task, the rewards and penalties incurred by the subject depend only on the motion trajectory that actually occurs. In our experiments, the penalties are explicit monetary rewards and penalties known to the subject. We will use the term "gain" to refer to both rewards and penalties, for convenience. Gains are associated with particular regions of space, denoted R_i, $i = 0,...,N$. If the subject's actual trajectory enters region R_i then the subject incurs a gain (reward or penalty) G_i. We refer to regions where the gain is positive as target regions or reward regions. We refer to regions where the gain is negative as penalty regions (a negative gain is a loss).

5. The expected gain function $\Gamma(S)$ of the motor strategy S is

$$\Gamma(S) = \sum_{i=0}^{N} G_i P(R_i|S) + G_{\text{timeout}} P(\text{timeout}|S) + \lambda B(S). \tag{1}$$

$P(R_i|S)$ is the probability, given a particular choice S of mean movement end point, of reaching region R_i before the time limit t = timeout has expired,

$$P(R_i|S) = \int_{R_{i,\text{timeout}}} P(\tau|S)\mathrm{d}\tau, \tag{2}$$

where $R_{i,\text{timeout}}$ is the set of trajectories that pass through R_i at some point in time after the start of the execution of the visual–motor strategy and before time t = timeout.

6. An optimal visual–motor strategy S on any trial is one that maximizes the subject's expected gain $\Gamma(S)$. The prediction of our model is that subjects will trade off risks entailed by different strategies so as to maximize their expected gain.

Biomechanical costs associated with the selected movement trajectory are included in the expected gain function in Eq. (1) by a biomechanical gain function $B(S)$, which is typically negative (a cost or penalty associated with the trajectory). As the tasks involve a time limit for response and a penalty for failure to respond before the limit, Eq. (1) contains a term reflecting this "timeout" penalty. The probability that a task leads to a timeout is denoted by $P(\text{timeout}|S)$ and the associated gain by G_{timeout}. The parameter λ in Eq. (1) characterizes the trade-off between physical effort and expected reward that the subject will tolerate.

Selecting the optimal motor strategy S corresponds to solving Eq. (1). Note that a complete solution of Eq. (1) is no trivial task: Any given motor strategy S implies a set of motor commands that, depending on the noise associated with the execution of these commands,[22] determines the motor variability associated with this strategy $[P(\tau|S)]$. The selection of $P(\tau|S)$, on the other hand, affects the probability of incurring rewards or penalties $[P(R_i|S)]$ or the timeout $[P(\text{timeout}|S)]$. Therefore finding the optimal motor strategy S in Eq. (1) corresponds to simultaneously computing $P(\tau|S)$ and optimizing Eq. (1).

In the following we will discuss predictions of our model in an experimental situation in which $P(\tau|S)$ [and $B(S)$] remain constant over the range of relevant motor strategies. In addition, in our experiment $P(\tau|S)$ can be estimated experimentally. This reduces the problem of simultaneously computing $P(\tau|S)$ and optimizing Eq. (1) to the much simpler problem of measuring $P(\tau|S)$ and using its estimate to solve Eq. (1).

3. EXAMPLE: SELECTING THE OPTIMAL MOVEMENT END POINT IN A LANDSCAPE OF EXPECTED GAINS

Imagine a subject carrying out the simple task discussed above. The subject is asked to touch within a target circle drawn on a computer screen. If the subject hits the target on a trial, he or she will win 100 points. After the experiment is over, the total point winnings are converted to a monetary reward. On all trials, a second "penalty" circle appears to the left of the target, partially overlapping it. If the subject hits within the penalty circle, she or he will lose 100 points. Note that as the regions overlap, the subject may hit within both of them and simultaneously incur the specified reward and the specified penalty, receiving zero points [see Fig. 2(b) for a display of the stimulus configuration]. We force the subject to touch the screen within a specified time limit that is the same on every trial. Late responses incur a very high penalty.

Each trial has three possible outcomes:

1. The target is hit: The subject receives a reward of G_0 points (given that we represent values as gains, $G_0 > 0$).
2. The penalty region is hit: The subject receives a penalty of G_1 points ($G_1 < 0$).
3. Time out: The subject was too slow and receives a large penalty of $G_{\text{timeout}} < 0$.

The first and second options are not mutually exclusive if the target and penalty regions overlap.

How does one predict the optimal movement end point that maximizes the gain function [Eq. (1)]? Given our effectively two-dimensional task, we will designate visual–motor strategies S by their resulting mean point of impact on the computer screen (x, y), i.e., the mean movement end point of the subject. For mean movement end point (x, y), the expected gain is

$$\Gamma(x, y) = G_0 P(R_0|x, y) + G_1 P(R_1|x, y) + G_{\text{timeout}} P(\text{timeout}|x, y) + \lambda B(x, y), \tag{3}$$

where G_0 is the gain associated with the reward region R_0, and G_1 is the penalty associated with the penalty region R_1.

In our experiments the probability of a timeout and the biomechanical gains are effectively constant over the limited range of relevant screen locations. In evaluating whether subjects correctly maximize expected gains, we can ignore the constant timeout and biomechanical penalty terms. Thus the subject must choose (x, y) so as to maximize

$$\Gamma'(x, y) = G_0 P(R_0|x, y) + G_1 P(R_1|x, y). \tag{4}$$

To compute the probabilities $P(R_i|x, y)$, we need to model the subject's motor uncertainty. We assume that pointing trajectories are unbiased and distributed around the mean movement end point (x, y) according to a Gaussian distribution,

$$p(x', y'|x, y) = \frac{1}{2\pi\sigma^2} \exp\{-[(x' - x)^2 + (y' - y)^2]/2\sigma^2\}, \tag{5}$$

where σ indicates the spatial motor variability of the subject's responses. We assume that the subject's motor variability is the same in the vertical and horizontal directions, on the basis of previous experience with subjects in similar tasks,[1] and we will test whether this assumption holds in the experiments reported below.

The probability of hitting region R_i is, then,

$$P(R_i|x, y) = \int_{R_i} p(x', y'|x, y)\mathrm{d}x'\mathrm{d}y'. \tag{6}$$

For the stimulus configurations used in our experiments (see Figs. 3 and 8, below) no analytical solution could be found for Eq. (6). The integral was solved by integrating Eq. (6) numerically,[29] and the results were used for maximizing Eq. 4.

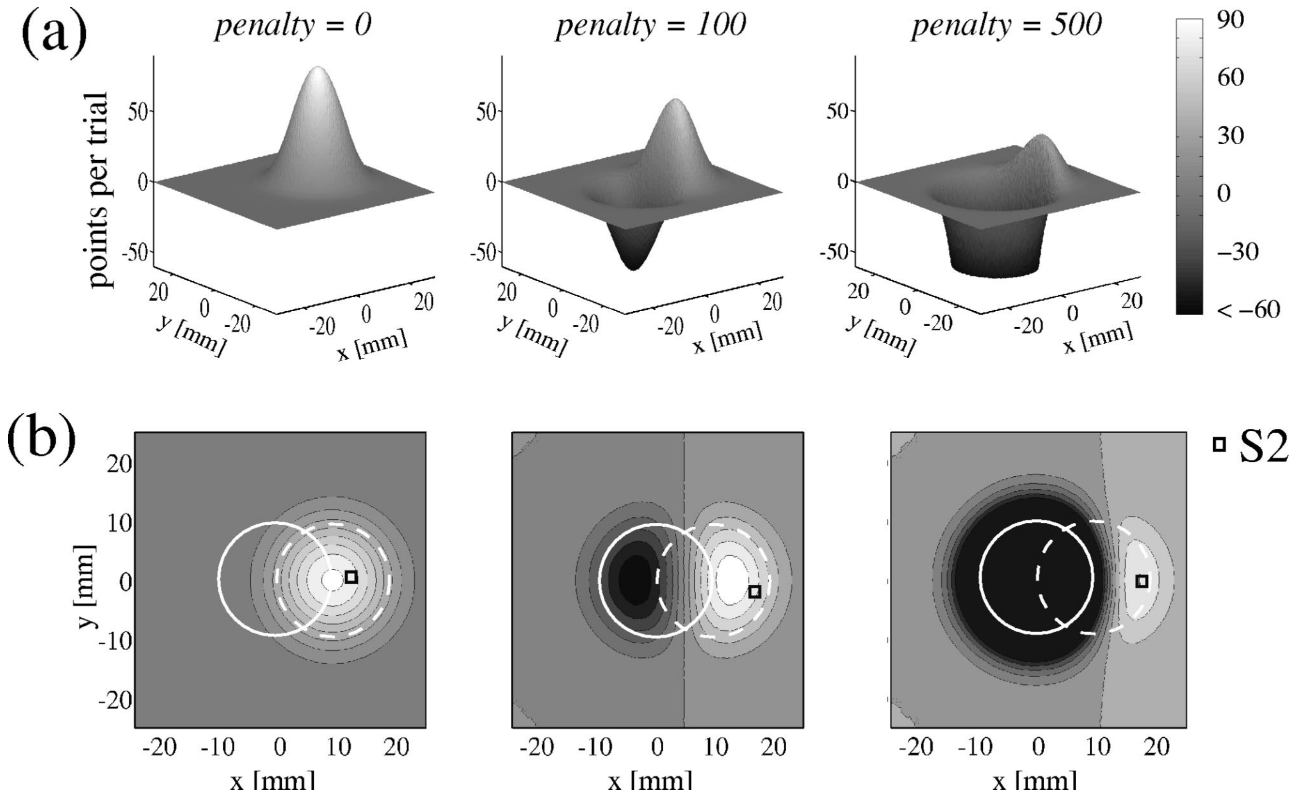

Fig. 2. "Landscape" of expected gain for an optimal observer with a variance of $\sigma^2 = 4.83$ (matching that of subject S2 in experiment 1). (a) Expected gain (in points per trial) as a function of the mean movement end point (x, y). The distribution is truncated for scores <-60 points. (b) The same landscape replotted as a contour plot with the mean movement end point of subject S2 (open squares) compared with optimal performance as predicted by the model [the contour regions are coded with the same gray-level scale as in (a)].

We first consider the gain G_1 associated with hitting the penalty region for a given mean movement end point. Suppose that G_1 is 0, i.e., there is no penalty associated with hitting the penalty region. Then we can reasonably expect that the subject will seek to hit the target as often as possible, given the time limit. The optimal mean movement end point is the center of the target under these conditions (Fig. 2(b), left panel). The subject's winnings depend only on his or her motor uncertainty, captured by the parameter σ in the specification of the Gaussian. This is illustrated in Fig. 2(a), which shows the winnings as a function of mean movement end point (i.e., the figure shows expected gains, which the observer should maximize). The left panel displays the case $G_1 = 0$. The largest gain is predicted if the observer's mean movement end point coincides with the target center. Note that owing to the subject's motor variability (equal to subject S2's motor variability in experiment 1; see Table 1 below) the subject occasionally misses the target, resulting in an average win of 82.1 points (out of a possible 100).

Next, suppose that hitting the penalty circle incurs a penalty of 100 points ($G_1 = -100$). Given the spatial motor variability of the pointing responses, the subject may accidentally hit the penalty circle for a mean movement end point at the center of the target. When trying to maximize the gain across all trials, it may be preferable to shift movement end points to the right of the center of the target. Depending on the amount of the penalty, it might be less costly to miss the target once in a while to avoid the risk of incurring so many penalties. Therefore we expect the optimal mean movement end point to shift farther to the right with increasing penalty (for a constant target area). This is demonstrated in Fig. 2(b) (middle panel) in which the mean movement end point resulting in maximal expected gain has shifted to the right. Note that the maximum expected gain in this condition (57.6 points) is less than that of the penalty = 0 condition. If the penalty value increases to 500 points, the optimal mean movement end point moves even further to the right [Fig. 2(b), right panel]. This means that the subject will safely avoid hitting the red penalty region but also will rarely collect points for hitting the green circle. The maximum expected gain drops to 30.2 points (Fig. 2, top-right panel).

We next describe two experiments designed to test the basic assumptions of our theory. The results of the experiments will be compared with the predictions of the model. We will estimate subjects' motor variabilities from the pointing data. We use that estimate of the motor variability to compute subjects' optimal mean movement end points for various stimulus configurations and penalties, using the model developed above. We compare these predictions with subjects' performance.

4. EXPERIMENT 1

In a previous experiment[1] we demonstrated that subjects modify their mean movement end points in response to changes in the gains associated with the penalty region. All the subjects' winnings were within 8% of that predicted by the model. The subjects' distributions of movement end points, however, were too noisy to allow a precise comparison with the spatial predictions of the model. We therefore decided first to measure subjects' performance using the same stimulus configurations but collecting sufficient data per subject to allow for a separate experimental test of our model for each subject. We also changed the time course of trials as described below.

A. Method

1. Apparatus

The experimental setup was the same as used previously.[1] Subjects were seated in a dimly lit room in front of a transparent touch screen (AccuTouch from Elo TouchSystems, accuracy $< \pm 2$ mm standard deviation, resolution of 15,500 touchpoints/cm^2) mounted vertically in front of a 21-in. computer monitor (Sony Multiscan CPD-G500, 1280×1024 pixels @75 Hz). A chin rest was used to control the viewing distance, which was 29 cm in front of the touch screen. The computer keyboard was mounted on the table centered in front of the monitor. The experiment was run with the Psychophysics Toolbox[30,31] on a Pentium III Dell Precision workstation. A calibration procedure was performed before each session to ensure that the touch screen measurements were geometrically aligned with the visual stimuli.

2. Stimuli

Each stimulus configuration consisted of a target region and a penalty region. The penalty region was circular, with light red shading, a bright red edge, and a small bright red circle marking the center. The target region was also circular, marked by a bright green edge, and unshaded so that the overlap with the penalty circle would be readily visible. The target and penalty regions had radii of 32 pixels/9 mm.

The target region was displaced horizontally from the penalty region, either to the left or right. There were three possible magnitudes of displacement. The six resulting stimulus configurations are summarized in Fig. 3. The penalty and target regions always appeared simultaneously, on a black background.

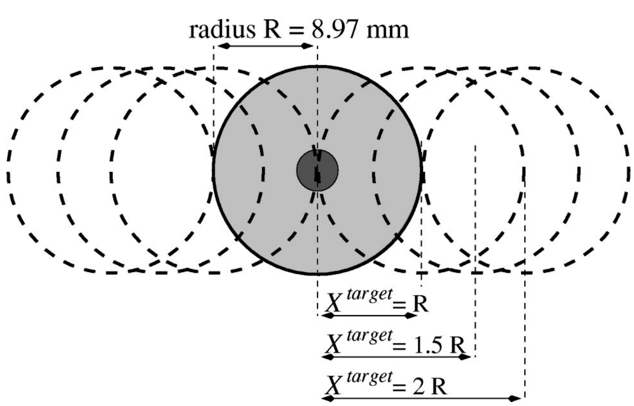

Fig. 3. Layout of the stimuli in experiment 1. The six dashed regions indicate the six different positions at which the target could appear.

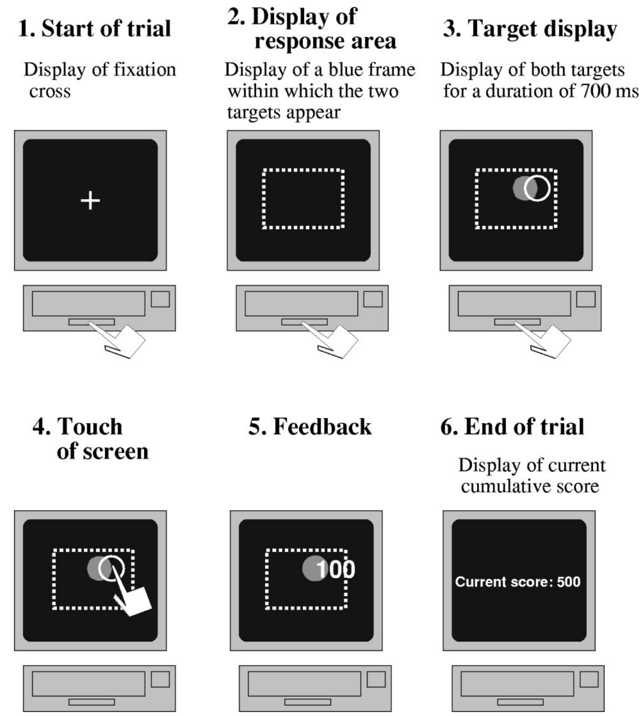

Fig. 4. Sequence of events during a single trial. The trial would not start until the subject depressed and held the space bar. In screen 5 (feedback) the subject would be shown each of the regions that she or he hit (the region would "explode" graphically) and the associated gain (penalty or reward) incurred.

To prevent subjects from using preplanned movements, the whole stimulus configuration was "jittered" by a randomly chosen amount in each trial; the shifts in x and y were chosen independently from a uniform distribution over the range ± 44 mm. A blue frame (114.2 mm $\times$ 80.6 mm), centered around the screen center, indicated the area within which the target and penalty regions could appear.

3. Procedure

Each trial followed the procedure illustrated in Fig. 4. A fixation cross indicated the start of the trial. The subject was required to depress the space bar of the keyboard with the same finger that she or he would later use to touch the screen. The trial would not begin until the space bar was depressed; the subject was required to hold the space bar down until after the green target appeared. Next, the blue frame was displayed, delimiting the area within which the targets would appear and preparing the subject to move shortly. After an interval of 500 ms the green target and the red penalty region were displayed simultaneously on the screen. Only after the appearance of the target was the subject allowed to release the space bar and initiate a movement toward the touch screen (allowing us to measure the time of movement initiation). After the green target was displayed, subjects were required to touch the screen within 700 ms or they would incur a "timeout" penalty of 700 points. If the subject touched the screen within the area of the red or the green target, the target(s) that were hit "exploded" graphically. Then the points awarded for that trial were shown, followed by the subject's total accumulated points for that session. A hit on the green target gained 100 points. The penalty for touching the red penalty region was constant during a block of trials and could be 0, 100, or 500 points (gains of 0, -100, -500). If the screen was touched in the region of overlap between the target and the penalty region, then the reward and the penalty were both awarded. If a subject anticipated the target display, releasing the space bar before or within 100 ms after display of the green target, the trial was abandoned and repeated later during the block.

Each block of trials consisted of five repetitions of each of the six target locations, presented in random order. The first session served as a practice session during which the subject learned the speeded motor task. In this session, the subjects first ran a single block in the penalty zero condition without a time limit for the response. This was followed by three blocks of trials with a moderate time limit of 850 ms, again followed by six blocks with a time limit of 700 ms. The practice session was followed by two single experimental sessions on a different day within the same week. Experimental sessions consisted of a touch screen calibration, 12 practice (warm-up) trials (two repetitions of the six target locations in the penalty $= 0$ condition), and 12 blocks of 30 trials (four blocks for each of the three penalty levels), presented in random order.

4. Subjects and Instructions

Five subjects participated in the experiment. The subjects were two male and three female members of the Department of Psychology at New York University. All participants were right-handed, had normal or corrected-to-normal vision, and ranged in age from 26 to 32 years. All subjects had given their informed consent before testing and were paid for their participation. All were unaware of the hypotheses under test. Subjects ran one practice session of 300 trials and two sessions of 372 trials (12 practice trials, 12 blocks of 30 trials) which took approximately 45 min per session. Subjects were informed of the payoffs and penalties before each block of trials. All used their right index finger to depress the space bar at the start of a trial and to touch the touch screen. Subjects were told that the overall score over the three sessions would be converted into a bonus payment at the rate of 25 cents per 1000 points so as to motivate fast, accurate responses.

5. Data Analysis

For each trial we recorded the reaction time (the interval from stimulus display to the release of the space bar), the movement time (the interval from release of the space bar to touching the screen), the response time (the sum of reaction and movement time), the screen position that was hit, and the score. Trials in which the subject released the space bar less than 100 ms after display of the green target or hit the screen more than 700 ms after display of the green target (timeouts) were excluded from the analysis.

Penalty coordinates. Each subject contributed approx. 720 data points, i.e., 40 repetitions per condition. Movement end-point positions (X_{ij}^p, Y_{ij}^p) were recorded relative to the center of the red penalty region (Fig. 3) for each penalty-value condition i ($i = 1, 2, 3$), displacement condition j ($j = 1,..., 6$), and trial p ($p = 1,..., n_{ij}$). In these coordinates, the six possible positions of the green target were $X_j^{\text{target}} = -17.9\ -13.4, -9, 9, 13.4$, and 17.9 mm and $Y_j^{\text{target}} = 0$ ($j = 1,..., 6$). We first computed the mean end point for each subject and each condition $\bar{X}_{ij}$ and $\bar{Y}_{ij}$ by averaging across replications $p = 1,..., n_{ij}$. We omitted those trials (31 out of 3600) where the subject did not reach the touch screen within the time limit (timeout).

Variance. For each subject, we tested whether variances in the x and y directions were independent of conditions and found that they were ($F < 1.35$ for all subjects). We found no evidence of correlation between the x and y directions, and, accordingly, we estimated a pooled variance σ^2 across the X and Y coordinates and all conditions for each observer separately.

Equality of reaction times, movement times, and Y coordinate of movement end points. Reaction times, movement times, and the Y coordinate of the movement end points were analyzed individually for each subject as a 1-factor, repeated-measures analysis of variance (ANOVA) across all 18 conditions (three penalty levels times six target locations).

Overall bias. We next analyzed whether each subject had a consistent overall bias. This bias represents a tendency to consistently aim to the left or right or up or down independent of condition. Because of the vertical and horizontal symmetry of the 18 conditions of the experiment, the horizontal bias X_{bias} is readily estimated by averaging the 18 values of $\bar{X}_{ij}$, and the vertical bias Y_{bias} is estimated similarly. Given the symmetries of the stimulus configurations, an evident prediction of the model is that $X_{\text{bias}} = Y_{\text{bias}} = 0$. We will test this prediction in the following section.

MEG predictions. Once we have an estimate of a subject's movement variability, σ^2, we can use the MEGaMove model to predict the maximum expected gain (MEG) end point (X_{ij}^{MEG}, Y_{ij}^{MEG}) that corresponds to maximum expected gain for that subject in that condition.

In conditions where the penalty associated with the penalty region $G_1 = 0$, the MEG point (X_{ij}^{MEG}, Y_{ij}^{MEG}) should be the center of the target region, (X_j^{target}, Y_j^{target}). That is, when there is no penalty the subject maximizes expected gain by ignoring the penalty region and selecting a distribution of movement end points with the mean at the center of the target region. This is a consequence of the symmetry of the error model that we employ.

Target coordinates. In comparing a subject's performance with optimal performance, it will be convenient to replot points relative to the center of the green target region. We define, $\bar{x}_{ij} = \bar{X}_{ij} - X_j^{\text{target}}$ and $\bar{y}_{ij} = \bar{Y}_{ij} - Y_j^{\text{target}}$ for each subject and all conditions. We similarly define (x_{ij}^{MEG}, y_{ij}^{MEG}) as the MEG point, now expressed as a shift away from the center of the green target region.

Score. To test how the subjects' winnings compared with optimal performance predicted by the model, the mean and variance of the distribution of optimal performance were computed. This was done by performing computer simulations of the experiment for each subject individually, i.e., simulating 40 trials in each condition by using the estimate of the subject's motor variance and computing the respective score in each experiment. One hundred thousand repetitions of the simulated experiment provided a precise estimate of the mean and the variance of the distribution of optimal performances. This estimate was used to test whether the subjects' performance was significantly different from optimal.

We will use both penalty (X, Y) and target (x, y) coordinate systems in the analysis below.

B. Results

As subjects differed significantly in their motor variability, reaction, and movement times, the data were analyzed individually for each subject. Results are reported in Table 1.

Equality of reaction and movement times. The results of the statistical analysis for each subject confirmed that the reaction and movement times did not differ significantly across conditions ($p > 0.05$ in all cases).

Response time distributions. In choosing their movement and reactions times, our subjects chose to use all of the time available to them consistent with avoiding all timeouts. The time limit of 700 ms fell at roughly the 99.9th percentile of the distribution of overall response times. Any shift of this distribution to longer response times would have increased the timeout rate, whereas a shift to faster responses times could only reduce it by 0.1%. This outcome is consistent with models that predict that subjects minimize movement variance by using all of the time available to move.[20,22,24]

Overall bias. Subjects showed a small constant rightward shift X_{bias} of movement end points ranging from 0.1 (subject S5) to 2 mm (subject S2). Those for subjects S1, S2, S3, and S4, were significantly different from 0 at the 0.05 level (t-test, $p < 0.05$, Bonferroni correction for five tests). Some subjects also showed small individual upward displacements Y_{bias} ranging from 0.3 ± 0.2 mm (subject S2) to 2.1 ± 0.2 mm (subject S1). Those for subjects S1, S3, S4, and S5 were significantly different from 0 at the 0.05 level (t-test, $p < 0.05$, Bonferroni correction for five tests). In carrying out the remainder of the analysis, we will report results for the data with and without correction for this bias, where appropriate, and return to a consideration of the biases below.

Effect of penalty and displacements: Y coordinates. Subjects' movement end points did not vary systematically across conditions in the y direction (data not reported here; contact authors for details). As just noted, we found a small but significant vertical bias for all subjects that was constant across all target positions and independent of penalty and displacement conditions.

Effect of penalty and displacements: X coordinates. Our model predicts a specific horizontal shift of the movement end point away from the red penalty regions for each choice of penalty and displacement. This shift should be stronger for target positions closer to the penalty region and for higher penalty values. We predict no shift in zero-penalty conditions, and in conditions with nonzero penalties the magnitude of shift should increase with increasing motor variability.

Figure 5 (left column) shows the raw mean x coordinates of the end points, $\bar{X}_{ij}$, for each target position, penalty condition, and subject in comparison with the respective expected optimal end points X_{ij}^{MEG}. As predicted by the model, the movement end points shifted away from the target center (target centers are indicated by the dotted vertical lines in Fig. 5). This shift was largest for the highest penalty value [Fig. 5(a)] and the target positions closest to the penalty region (represented in the figure by the points above the largest X_{ij}^{MEG} values). The small rightward bias X_{bias} mentioned earlier is visible in these plots as a small displacement upward of the data points relative to the 45-deg line.

In Figs. 5(a) and 5(b), right column, we collapsed over right–left symmetric conditions. (We tested whether the pattern of movement end points was symmetric, and for all subjects the pattern of movement end points did not differ significantly from symmetry; t-tests for each subject and penalty condition were $p > 0.05$ in all cases.) Data were replotted for the $G_1 = -100$ and $G_1 = -500$ penalty conditions (no shift predicted for $G_1 = 0$). We plotted for each subject the unsigned magnitudes of the horizontal shifts away from the center of the target (the absolute values of the target coordinates $\bar{x}_{ij}$) versus the corresponding unsigned magnitudes of the optimal shifts (the absolute value of x_{ij}^{MEG}). We removed the constant bias X_{bias} from the data to facilitate comparison of the effect of changes in displacement and penalty with the changes predicted by the MEGaMove model.

The 45-deg solid lines in Figs. 5(a) and 5(b), right column, indicates a perfect correspondence between the subjects' shift and the predictions of the model. Overall, subjects' results followed the predictions of the model. The results of the statistical analysis are consistent with these observations. As displayed in Fig. 5(a), right column, subjects with a higher motor variability (S2 and S4) shifted farther away from the target center than less-variable subjects (S1, S3, and S5), consistent with the predictions of the model.

We fitted separate regression lines to the data in Figs. 5(a) and 5(b), right column, for each subject and for the $G_1 = -100$ and $G_1 = -500$ penalty conditions. We performed two tests on the estimated slope for each subject and condition. We first tested whether the slope of this line is zero versus the alternative that it is greater than zero. A slope of zero indicates that the subject is not altering mean end point in response to changes in penalties or displacements. Not surprisingly, we rejected the null

Table 1. Experiment 1, Results[a]

Subject	σ^2 (mm^2)	Reaction Time	Movement Time	X_{bias} (mm)	Score ($)
S1	11.94	179 ± 34 ms	389 ± 28 ms	1.0 ± 0.1	15.73
S2	23.36	213 ± 44 ms	344 ± 39 ms	2.2 ± 0.2	13.08
S3	11.07	205 ± 37 ms	381 ± 37 ms	1.0 ± 0.1	15.80
S4	19.65	253 ± 40 ms	285 ± 40 ms	1.7 ± 0.1	14.58
S5	11.43	215 ± 43 ms	352 ± 31 ms	0.1 ± 0.1	15.90

[a] Data reported for the five subjects individually; spatial motor variability (σ^2), reaction and movement times (± one standard deviation), and constant response bias in the x direction X_{bias} (± one standard error of the mean) computed by averaging across all conditions (~720 data points per subject); the score indicates the cumulative sum of wins across all trials.

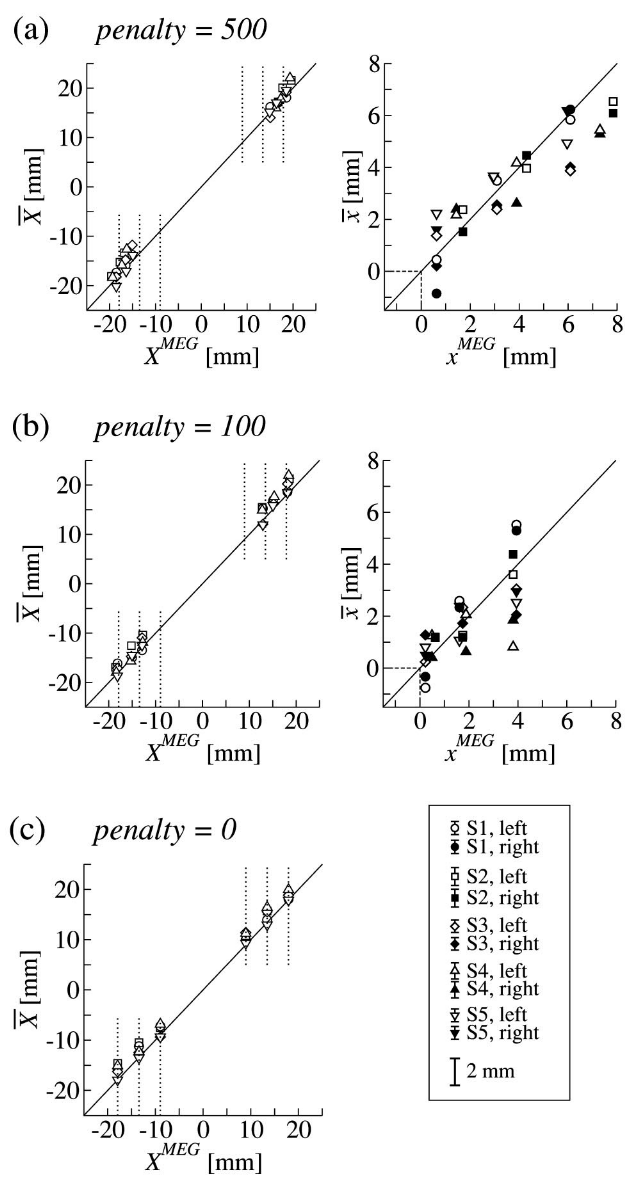

Fig. 5. Experiment 1, results for five subjects under penalty conditions 0, 100, and 500. Left column, mean movement end points $\bar{X}_{ij}$ (x coordinate) as a function of the optimal mean movement end point X_{ij}^{MEG} predicted by the model for five subjects. Model predictions based on each subject's variability were computed. Solid lines, perfect correspondence of model and experiment; dotted lines, center of the green target region. Data are uncorrected for bias; pointing bias is visible as a shift of the data above the prediction line. Right, shift of mean movement end points from the center of the green target region, corrected for pointing bias ($\bar{x}_{ij}$) as a function of the optimal shift of mean movement end point (x_{ij}^{MEG}). Data for green target regions to the *left* of the penalty region (open symbols) were reflected; solid symbols indicate mean movement end points toward green target regions on the right of the penalty region. Data were corrected for constant pointing bias by subtracting the constant pointing bias given in Table 1. Average standard error of the mean is indicated in the key. Right, solid symbols indicate data for targets on the right side of the configuration.

hypothesis for both of the penalty conditions, $G_1 = -100$ and the $G_1 = -500$ ($p < 0.0001$ for all subjects and conditions).

Next, we tested the hypothesis that the slope is precisely 1 versus the alternative that it is not. A slope of 1 implies that the subject's shifts in response to changes in penalty and displacement are precisely those predicted by the model. A slope less than 1 (but greater than 0) indicates that the subject shifted mean end point less than is optimal in response to changes in penalties and displacements. A slope greater than 1 indicates that the subject shifted mean end point more than is optimal in response to changes in penalties and displacements. That is, the subject was overly "afraid" of the red penalty.

Table 2 contains the slope estimates and raw p values for the null-hypothesis test that the slope is 1 (employing a Bonferroni correction for ten tests, significant deviations from the model predictions with $p < 0.05$ marked by an asterisk). We rejected the null hypothesis in both penalty conditions for two of the five subjects (S3 and S4) in both penalty conditions. The slope estimates for these subjects in both conditions were less than 1. This indicates that these two subjects did not shift away from the red penalty region as far as they should have to optimize maximum expected gain.

Subject S1's movement end points deviated significantly from the model predictions in the penalty = 100 condition, but in the opposite direction. The slope of 1.38 indicated that S1 overreacted to changes in the relative location of the red penalty region more than an optimal mover should.

At least some of the subjects did not fully maximize expected gain as predicted by the model. Of course, if we collected sufficient data in each condition, we would eventually expect to be able to discriminate any human subject's performance from optimal performance. No physical coin is ever really fair, and no human observer is ever really ideal. It is therefore important to assess the consequences of the subjects' deviations from the predictions of the model in terms of a second standard, the amount of money won or lost over the course of the experiment. Note that subjects were not instructed to aim as accurately as possible (they were not even told where to aim). Rather, they were told to respond so as to maximize their payment, something they were likely motivated to do in any case.

Therefore we compared the subjects' winnings for each target position and penalty condition with the optimal possible score as predicted by the model. In the penalty = 0 condition this "optimal" score is determined by the

Table 2. Experiment 1, Comparison of Results with Model Predictions[a]

Subject	Penalty = 100 Slope	Penalty = 100 p	Penalty = 500 Slope	Penalty = 500 p	Performance (%)
S1	1.38±0.08*	<0.0001	0.97 ± 0.05	0.5824	98.60
S2	1.03±0.11	0.7740	0.86 ± 0.05	0.0067	104.92
S3	0.73±0.08*	<0.0001	0.68 ± 0.05*	<0.0001	97.57
S4	0.76±0.08*	0.0003	0.74 ± 0.04*	<0.0001	107.67
S5	0.86±0.09	0.1298	1.00 ± 0.05	0.8886	98.91

[a] Model predictions were computed for each of the five subjects individually by using estimates of individual subjects' motor variability σ^2 as given in Table 1. See text for details.

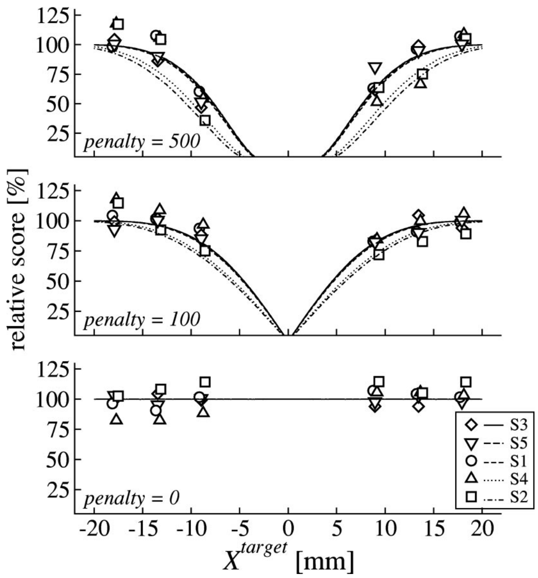

Fig. 6. Experiment 1, results for five subjects, listed in order of motor variability, for penalty conditions 0, 100, and 500. The values plotted on the vertical axis are average scores per target position displayed as a percentage of optimal performance predicted by the model for penalty = 0. Normalizing in this way makes it easier to compare performance of subjects with different motor variabilities. The horizontal axis is the target position X^{target} relative to the penalty region. Model predictions based on each subject's variability were computed. The curves (one per subject) represent the model predictions.

subject's motor variability: A more variable subject will miss the target more often, winning fewer points on average. To compare the performance of individual subjects, we therefore normalized the individual data by dividing their wins per trials by the optimal score at penalty $= 0$. The resulting data provide an estimate of the degree to which each subject outperforms or underperforms relative to the optimal performer and shows how their winnings deteriorate with increasing proximity to the penalty region and increasing penalty (see also Fig. 2). The data of all five subjects are displayed in Fig. 6 and in Table 2 in which the total winnings are compared with optimal performance. In accordance with the model, subjects with lower motor variability (S1, S3, S5) scored higher than more variable subjects (S2, S4).

However, subjects' performance did not differ significantly from the predicted optimal performance (Table 2)—not even for subjects S1, S3, and S4 whose movement end points had been found to be significantly different from the model predictions. We therefore conclude that our model compares well with the experimental data.

Overall bias. We found constant biases across conditions in both horizontal and vertical directions for all subjects. These biases ranged from 0.3 to 2.1 mm in the *x* direction and 0.1 to 2.2 mm in the *y* direction. This bias is considerably smaller than the width of the finger tip. Note also that subjects never received experimental feedback on where they had hit the screen other than an exploding target circle (or the lack thereof when they missed) and that the subject's finger tip occluded the point of impact on the screen.

We suspect that these biases are due to a calibration failure. Subjects carefully aligned the touch screen with the visual stimuli before each experimental session by pointing at calibration points on the screen using *slow*, deliberate pointing movements. The location registered by the touch screen is the centroid of the part of the screen where the finger pad exerted more than a threshold pressure. During the experiment, the subjects moved rapidly. The part of the finger pad that impacted the screen could not be controlled, and the distribution of pressure was almost certainly different. Consequently, we hesitated to assign much importance to these small deviations.

To test this calibration-bias hypothesis, subject S1 repeated the entire experiment using the index finger of the *left* hand. After a practice session, S1 was able to perform the task without difficulty. Winning a bonus of \$15.63, S1 scored as high as with the right hand (Table 1). The motor variability ($\sigma^2 = 10.6\,\text{mm}^2$) as well as the response and movement times (177 ± 26 ms and 410 ± 27 ms) were equivalent to those previously measured when S1 had been responding with her right hand. In contrast to the previous outcome, movement end points deviated to the *left* (average deviation from the model predictions -0.75 ± 0.1 mm). We tentatively conclude that the consistent rightward deviation of the subjects' movement end points from the model predictions is a problem with calibration. Of course, we cannot use speeded movements in calibration, since subjects cannot accurately control the point of contact with the screen when moving rapidly. In the analysis, we used data corrected for these biases, except as noted.[32]

Reinforcement-learning models. Our results suggest that the subjects selected their movements in good agreement with the predictions of the MEGaMove model. Of course, we cannot conclude that they are selecting motor strategies by solving Eqs. (4)–(6) (Section 3).

How do subjects achieve near-optimal performance (as they do, at least in the experiment just reported)? A first class of models, of which the MEGaMove model is representative, consist of the assumptions that (1) the subject's motor system contains some representation of the subject's movement uncertainty and that (2) given information about the rewards and penalties associated with movement in a scene, the subject's motor system can effectively select the motor strategy that maximizes expected gain.

As we noted in Section 1, there is evidence that human movement planning takes into account movement uncertainty, and the results of experiment confirm this claim. There is a vast literature on cognitive decision making indicating that subjects are sensitive to gains and probabilities in deciding between simple gambles.[33] The issue, then, is whether the motor system employs the same or similar principles in deciding among an indefinitely large number of motor strategies.

Consider then an alternative explanation of subjects' performance in experiment 1. Subjects use a simple reinforcement-learning algorithm to find the motor strategy that maximizes expected gain. They search the reward landscape of Fig. 2 until they find the maximum. The observer, for example, may choose a motor strategy at random and then integrate the rewards received over many trials to estimate the expected gain. The subject

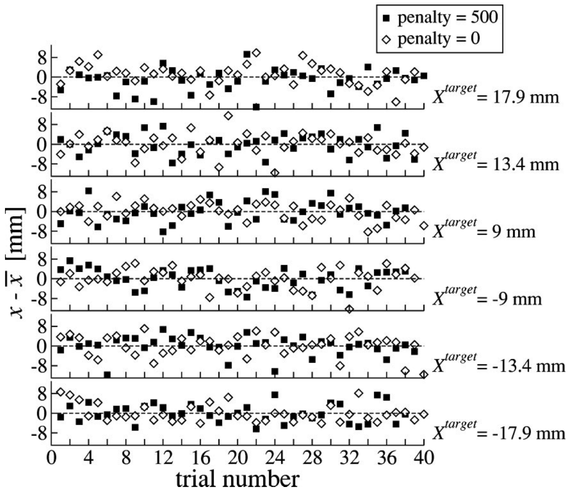

Fig. 7. Experiment 1, trial-by-trial data of subject S1. Deviations of individual movement end points (x coordinate) from the mean movement end point $\bar{x}$ (per condition) as a function of trial number are shown. Data are displayed for each location of the green target region and for two penalty conditions.

could then change to a second motor strategy and evaluate it, using the information gained to search the space of motor strategies seeking to maximize expected gain.

Suppose that we examined the performance across time of a "hill-climbing" movement planner. Initially, end points need not be near the end point associated with maximum expected gain, but we would expect a shift toward X^{MEG} over the course of any block. To determine the optimal mean movement end point, any reinforcement strategy would require trials in which the penalty region was hit. In fact, subjects hit the penalty region extremely rarely in this experiment: on fewer than 1 trial in 100 (in the penalty = 500 condition).

Furthermore, the shift toward the optimal end point is based on collecting information about the expected gain associated with a series of motor strategies. Yet the outcome of any choice of motor strategy is stochastic, and the observer must integrate over several attempts using one motor strategy to get a useful estimate of its expected gain. Hence, if subjects relied on a reinforcement-learning algorithm, we would expect to see a slow shift in end point toward the optimal over the initial trials in each block.

In experiment 2 we will explicitly look at the ability of practiced subjects to adapt rapidly to novel experimental configurations. We can, however, look at the time course of results in experiment 1 to see whether there is any hint of the slow shift in end point associated with reinforcement-learning models.

As displayed in Fig. 7, this is not the case. Movement end points randomly fluctuate with constant variability around the mean within each condition over the course of the entire experiment. There is no discernible trend at the beginning of each block (i.e., for every sixth trial). In particular, we can compare performance in the penalty = 0 and penalty = 500 conditions. In the former, the optimal mean movement end point is obvious (i.e., the center of the green target), whereas in the latter, it is not. Yet the distributions of end points at the beginnings of blocks in both conditions appear to be very similar.

5. EXPERIMENT 2

The results of experiment 1 suggest that subjects select mean movement end points by rapidly solving an optimization procedure in each trial. Consequently, we expected that they would be able to perform the same pointing task with a rotated but otherwise unchanged stimulus configuration without any additional practice. We expected subjects to point at the unfamiliar, rotated configurations with the same low motor variability as in the previous experiments.[34]

We performed a second experiment to directly test the assumption that subjects are able to transfer previously gained knowledge about their own motor variability when pointing at novel stimulus configurations. Second, we increased the level of difficulty of the movement planning task to test whether subjects still managed to plan the movement end point according to our theory in more complex situations.

First, we rotated the previously horizontal stimulus configuration (Fig. 3) by ±90 and ±30 degrees (Fig. 8, upper row). Since the same subjects participated in experiment 2 as in experiment 1, they were well practiced in the task, and it is plausible that they had already gained knowledge of their own motor variability.

Second, we added four more conditions experiment 2 in which we increased the complexity of the stimulus configuration by increasing the number of penalty areas to two. As displayed in Fig. 8, we chose a set of four configurations that were combinations of the first four configurations.

A. Method

1. *Apparatus*

The experimental setup was the same as in experiment 1.

2. *Stimuli*

The stimulus configurations of the stimuli are similar to those of experiment 1. The red penalty region and the green target region appeared in eight different combinations. Figure 8 (upper row) shows configurations 1–4, composed of a single green target region and a single red penalty region. In configurations 5–8, a single green target region was combined with two red penalty regions, combining two of the first four configurations (Fig. 8, lower row). Target and penalty regions had radii of 32 pixels/9 mm. Red penalty areas were surrounded by a bright red edge, making the position of each penalty area

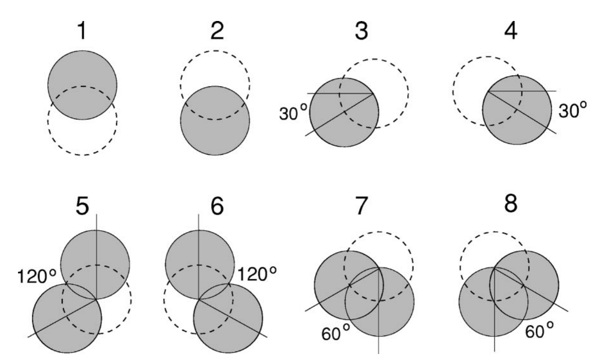

Fig. 8. Layout of the stimuli in experiment 2.

clear as well as indicating where the two penalty regions overlapped. The green target was a transparent overlay as before.

3. Procedure

The sequence of events during a single trial remained unchanged from experiment 1. Again, the subject could win 100 points by hitting the green target. The penalty for touching either of the red penalty regions was constant during a block of trials and varied randomly between blocks taking values of 0 or 500 points. Each block of trials consisted of four repetitions of each of the eight configurations presented in random order. As before, each session began with a touch screen calibration followed by 16 practice trials (each stimulus configuration displayed twice in random order for a penalty = 0 condition). Finally, 12 blocks (6 at each penalty level) of 32 trials were presented in random order.

4. Subjects and Instructions

The same five subjects participated as in experiment 1. Subjects ran one session of 400 trials (16 practice trials, 12 blocks of 32 trials), which took approximately 50 min. As before, subjects were told that the total score would result in a bonus payment of 25 cents per 1000 points.

5. Data Analysis

For each trial we recorded reaction time, movement time, the screen position that was hit, and the score. Trials in which the subject released the space bar less than 100 ms after display of the green target or hit the screen more than 700 ms after display of the green target were excluded from the analysis.

Each subject contributed approximately 384 data points, i.e., 24 repetitions per condition. Movement endpoint positions were measured relative to the center of the green target region. The green target was always presented at the screen center, i.e., $X^{\text{target}} = Y^{\text{target}} = 0$ for all configurations. The subject's motor variability was averaged across the eight conditions and the x and y directions.

As in the previous experiment, all subjects showed a systematic pointing bias in both the x and y directions. For each subject this bias was estimated as the movement end points averaged across all eight configurations in the penalty = 0 condition and is displayed in Table 3 (the displays were not symmetric with respect to the y direction, so we could not include the nonzero penalty conditions as we did for experiment 1). Compared with the experimental results, the model predictions were shifted by each subject's constant pointing bias X_{bias} and Y_{bias}.

An approach similar to that of experiment 1 was chosen to compare the recorded movement end points with the model predictions. But, unlike in experiment 1, shifts in predicted movement end points occurred along both spatial dimensions. Overall, the model accounted for approximately 70% of the variance of the average movement end points (Table 4). Since variances in the x and y directions did not differ significantly, data were analyzed with a polar coordinate system. For each recorded movement end point the absolute distance from the origin r_{hit} and the angle with the positive x axis ϕ_{hit} [$\phi_{\text{hit}} \in [0, 360)$] were computed. These estimates were compared with the predictions of the model (r_{opt}, ϕ_{opt}) by computing the slopes r_{hit} versus r_{opt} and ϕ_{hit} versus ϕ_{opt} (data combined across all spatial configurations). Again, both slope estimates were constrained by a constant of zero, because data had been corrected for constant pointing bias. The slope estimates were compared with a slope of 1, indicating perfect correspondence between the data and model. Table 4 contains the slope estimates and the corresponding p values (df ~ 192, varying according to the exact number of data recorded per subject, data that Bonferroni corrected for number of subjects; significant deviations from the model predictions with $p < 0.05$ marked by an asterisk; please contact authors for details).

Subjects' winnings were compared with optimal performance by using the same procedure as in experiment 1. Significant deviations from optimal performance are indicated by an asterisk in Table 4.

B. Results

The number of recorded movement end points included in the analysis was 1902; 18 responses were omitted for be-

Table 3. Experiment 2, Results[a]

Subject	σ^2 (mm^2)	X_{bias} (mm)	Y_{bias} (mm)	Score ($)
S1	7.84	2.4 ± 0.2	1.0 ± 0.2	6.05
S2	14.78	0.8 ± 0.3	−0.5 ± 0.3	5.13
S3	7.71	0.7 ± 0.3	0.1 ± 0.3	7.13
S4	8.47	2.2 ± 0.2	1.2 ± 0.3	5.93
S5	8.93	1.2 ± 0.2	1.1 ± 0.3	7.20

[a] Data reported for the five subjects individually; motor variability (σ^2), response bias in the x and y directions (X_{bias}, Y_{bias}, ± one standard error of the mean), computed by averaging across all conditions (~384 data points per subject); the score indicates the cumulative sum of wins across all trials. Performance is computed by dividing the score by the score of an optimal performer with the same motor variability.

Table 4. Experiment 2, Comparison of Results with Model Predictions[a]

Subject	R^2 (%)	Slope (ϕ_{hit})	p	Slope (r)	p	Performance (%)
S1	75.31	0.94 ± 0.02	0.0076	1.10 ± 0.02*	<0.0001	77.0*
S2	67.53	0.98 ± 0.02	0.3608	1.06 ± 0.03	0.0542	91.0*
S3	68.09	0.99 ± 0.02	0.6105	0.87 ± 0.03*	<0.0001	90.3*
S4	65.39	1.03 ± 0.02	0.1995	1.30 ± 0.03*	<0.0001	79.5*
S5	73.31	0.96 ± 0.02	0.0723	1.15 ± 0.03*	<0.0001	95.1

[a] Model predictions were computed for each of the five subjects individually by using estimates of individual subjects' motor variability σ^2 as given in Table 3.

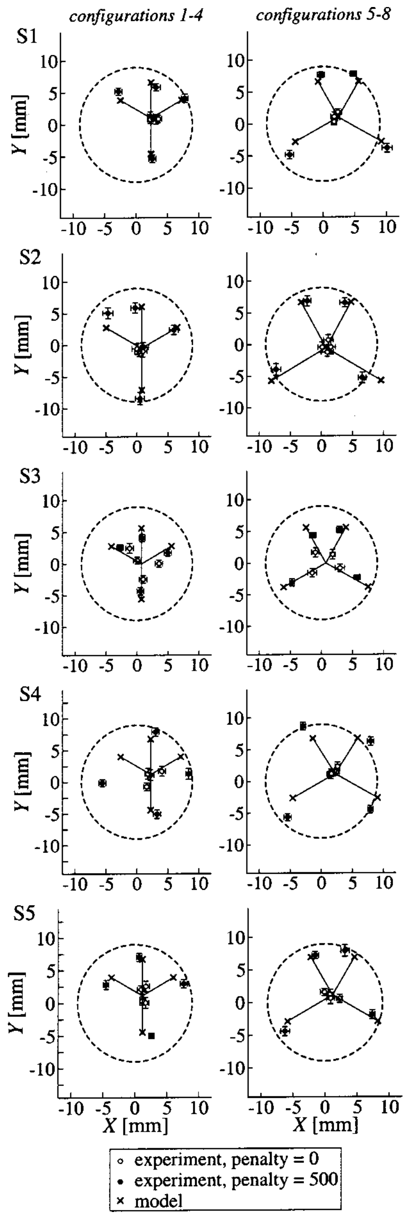

Fig. 9. Experiment 2, results: average movement end points ($\bar{X}$, $\bar{Y}$) for each of the eight stimulus configurations and the two penalty conditions, for each of the five subjects. Model predictions based on each subject's variability were computed. Left column, results for conditions 1–4 (one red penalty region). Right column, results for conditions 5 to 8 (two red penalty regions); see Fig. 8 for details. Error bars represent ± one standard error in the x and y directions, computed from 24 data points per condition (less than 24 in the rare cases where data were dropped owing to a timeout). Model predictions were corrected for each subject by adding the constant pointing bias given in Table 3.

ing late. As in the previous experiment, subjects differed significantly in their motor variability. Therefore the data were analyzed individually for each subject. Results are summarized in Tables 3 and 4 and displayed in Fig. 9. Reaction and movement times remained constant across conditions (data not reported here; contact authors for details), indicating that subjects did not alter their planning and movement strategy when pointing at configurations 5–8, i.e., for more complex stimulus configurations. There was no significant difference between the motor variabilities in the x and y directions nor across conditions. In particular, the distribution of responses was not affected by changes in the number of red penalty areas. Subjects did not reshape the distribution on the basis of the shape of the penalty region, nor did they engage in a speed–accuracy trade-off that varied between conditions.

A systematic pointing bias in both the x and y directions was apparent for all subjects. As in the previous experiment, this bias did not change with experimental conditions but differed between individual subjects (Table 3). We believe that the bias does not reflect systematic errors in the movement planning process but is a touchscreen bias due to the subject's use of the right hand (see the detailed discussion in the context of experiment 1, Subsection 4.B). We therefore corrected for the pointing bias in comparing the predictions of the MEGaMove model with the experimental data.

Figure 9 displays the mean movement end points of each of the five subjects for each stimulus configuration; it shows that, overall, movement end points follow the predictions of the statistical decision-theory model. The results of the statistical analysis support this observation. All subjects shifted their mean movement end points in the direction as predicted by the model (as reflected by slope estimates ~1 for ϕ_{hit}). The magnitude of this shift (indicated by r_{hit}) however, differed significantly from the model predictions. Three subjects (S1, S4, S5) shifted their mean movement end points farther away from the penalty region than predicted by the model, and one subject (S3) did not shift as far as predicted (Fig. 9 and Table 4).

In addition, four subjects (S1, S2, S3, S4) performed significantly below optimal performance predicted by the model. As is apparent from Fig. 10, subjects performed particularly poorly when pointing at configuration 6.

Two subjects reported that they had no idea why in particular in this condition they missed the green target, and even worse, hit the lower red penalty area, so frequently. Subject S1 hit the red penalty area four times in this condition (compared with only twice in all other seven conditions combined), S2 hit the red six times in this condition

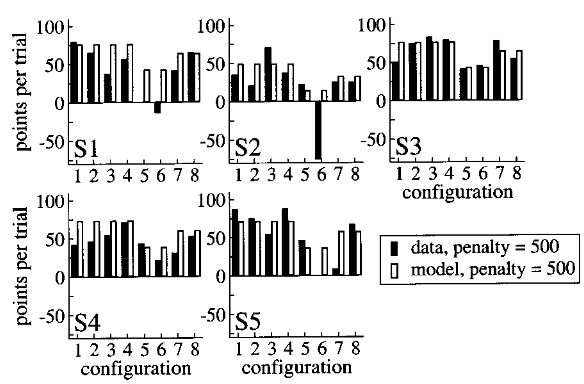

Fig. 10. Average gain per trial and model predictions in the penalty = 500 condition as function of the configuration (Fig. 8). Data displayed for each of the five subjects are based on 24 data points per configuration (less than 24 in the rare cases where data were dropped owing to a timeout). Model predictions based on each subject's variability were computed.

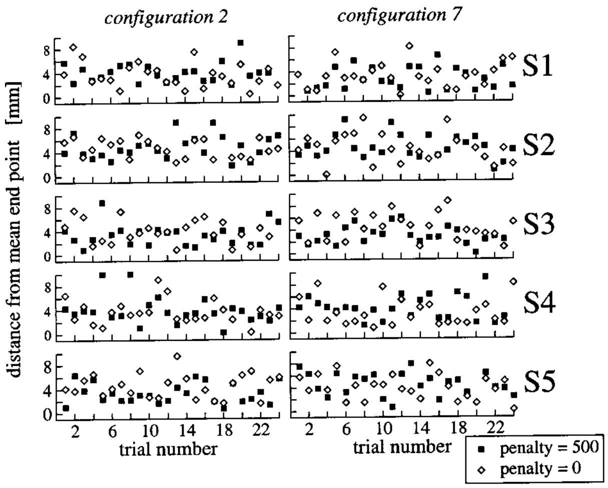

Fig. 11. Experiment 2, trial by trial data. Two-dimensional distance of individual movement end points (x, y) from the mean movement end point $(\bar{x}, \bar{y})$ as a function of trial number, displayed for configurations 2 and 7 (Fig. 8), for all five subjects individually and for penalty conditions 0 and 500.

(compared with five hits total in the other conditions), and S4 and S5 each hit the red three times (compared with three hits total in the other conditions). The reason for this might be the rightward pointing bias, which should lead to the most dramatic consequence (accidentally hitting red) when pointing at configuration 6 (Fig. 8). This is more severe for configuration 6 than for the configurations of experiment 1, as there is more penalty area in configuration 6 located just to the right of the green target. Only the performance of subject S3 deviated from optimal performance for a different reason: On average S3 did not shift her movement end points as far as predicted by the model, and as a consequence she collected penalties for hitting the red penalty across all conditions (1.3 hits per configuration).

In general, however, the MEGaMove model accounted for the subjects' behavior. The fact that subjects failed to perform close to optimally in this experiment does not indicate the use of suboptimal visual–motor strategies. Rather, it resulted from a nondetectable systematic response bias due to subjects' use of the right hand, which turned out to be costly in the context of this experiment. That subjects persisted in this particular suboptimal behavior also is a further indication that the basic correspondence of data and model is not the result of a reinforcement-learning procedure.

Finally, note that all subjects' motor variabilities were low and stable, as in the previous experiment, although the eight configurations constituted a novel set of stimuli and required planning motor responses in an unfamiliar context (pointing vertically and in tilted directions compared with the horizontal arrangements in the previous experiments). Subjects had only 16 practice trials (in a penalty = 0 condition) before the experiment began, and no learning effects were visible in the data across experimental blocks (Fig. 11). This suggests that during the previous experimental sessions subjects had gained knowledge about their own motor variability and were able to use this knowledge in a novel and more complex motor task. With this knowledge, the motor system is able to solve complex two-dimensional integration and optimization problems.

6. GENERAL DISCUSSION

We recently presented a model of movement planning based on statistical decision theory.[1] In this MEGaMove model, a motor strategy is selected by maximizing an expected gain function [Eq. (3)] that takes into account explicit gains associated with the possible outcomes of the movement, motor variability, biomechanical gains, and gains associated with time limits imposed on the mover.

In this study we performed two experiments to test the range of validity of our model. In these experiments subjects had to reach and hit a target region on a computer screen. Hitting the target within a prescribed time limit gained them a monetary reward. There were also one or more penalty regions on the screen that could partially overlap the target region. Hitting the regions incurred a specified penalty. In each experiment the penalty associated with hitting a penalty region as well as its position relative to the target region were varied. Using our model and an estimate of each subject's motor variability based on his or her performance, we could predict the mean movement end point that would maximize expected gain. Subjects followed the predictions of our model. With increasing penalty values, they shifted their mean movement end points away from the penalty region. This shift was larger for closer penalty regions and higher penalty values.

In the second experiment, subjects were exposed to a novel set of stimuli that consisted of four rotated configurations of experiment 1 and four more complex configurations consisting of one target and two penalty regions. We asked whether subjects still planned their movements as predicted by our model. Recall that the predictions of our model are based on the maximization of the expected gain function, which in the case of the four more-complex configurations requires a two-dimensional integration procedure. Subjects altered their movement end points in direction and magnitude as predicted by our model in all configurations presented. Furthermore, their motor variabilities were as low as in the previous experiments. Thus subjects were able to use their previously acquired estimate of motor variability in planning their responses in this more complex situation. And, the more complex array of penalties did not result in an increase in variability, owing to variability in the computation of the intended movement end point. This suggests that the bulk of the variability is due to execution of the motor command rather than to variability in the movement plan itself.

There are several things that subjects must know to calculate the optimal movement plan, including their motor variability (due to errors in both movement planning and movement execution), the target and penalty region positions and sizes, and the gains associated with these. The current experiments do not directly test whether each of these terms is estimated correctly by the subjects. However, although subjects initially required an hour of practice for variability to stabilize, variability remained stable with almost no practice trials across both experiments with different penalty and target conditions. Thus subjects must have acquired an estimate of their motor variability. Novel experimental conditions did not result in an increase in that variability, suggesting that movement planning for more-complex tasks (in experiment 2) did not result in additional variability.

Our model predicts optimal movement end points on the basis of a complex two-dimensional integration procedure. We therefore expect the model to fail once this integration procedure becomes too difficult or too time costly. It is likely that an increasing complexity of the stimulus configuration (containing a variety of different penalty regions or more-complicated stimulus shapes) and a reduced time limit on the task will push the limits of our model. Future experiments will study this question.

ACKNOWLEDGMENTS

We thank Geir Jordet for stimulating discussion and the reviewers for valuable comments on the manuscript. This work was supported by National Institutes of Health grant EY08266, Human Frontier Science Program grant RG0109/1999-B, and an Emmy–Noether fellowship (the Deutsche Forschungsgemeinschaft) to Julia Trommershäuser.

Corresponding author Julia Trommershäuser can be reached as follows: Department of Psychology, New York University, 6 Washington Place, 8th Floor, New York, N.Y. 10003; phone, 212-998-7853; fax, 212-995-4349; e-mail, trommer@cns.nyu.edu.

REFERENCES AND NOTES

1. J. Trommershäuser, L. T. Maloney, and M. S. Landy, "Statistical decision theory and trade-offs in the control of motor response," Spatial Vis. **16**(3–4), 255–275 (2003).
2. J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. (Springer, New York, 1985).
3. D. Blackwell and M. A. Girschick, *Theory of Games and Statistical Decisions* (Wiley, New York, 1954).
4. T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach* (Academic, New York, 1997).
5. L. T. Maloney, "Statistical decision theory and biological vision," in *Perception and the Physical World*, D. Heyer and R. Mausfeld, eds. (Wiley, New York, 2002), pp. 145–189.
6. T. Kaminsky and A. M. Gentile, "Joint control strategies and hand trajectories in multijoint pointing movements," J. Motor Behav. **18**, 261–278 (1986).
7. J. F. Soechting and F. Lacquaniti, "Invariant characteristics of a pointing movement in man," J. Neurosci. **1**, 710–720 (1981).
8. M. Dornay, Y. Uno, M. Kawato, and R. Suzuki, "Minimum muscle-tension chance trajectories predicted by using a 17-muscle model of the monkey's arm," J. Motor Behav. **2**, 83–100 (1996).
9. T. Flash and N. Hogan, "The coordination of arm movements: An experimentally confirmed mathematical model," J. Neurosci. **5**, 1688–1707 (1985).
10. Y. Uno, M. Kawato, and R. Suzuki, "Formation and control of optimal trajectory in human multijoint arm movement: minimum torque-change model," Biol. Cybern. **61**, 89–101 (1989).
11. J. F. Soechting, C. A. Bunea, U. Herrmann, and M. Flanders, "Moving effortlessly in three dimensions: Does Donders' law apply to arm movement?" J. Neurosci. **15**, 6271–6280 (1995).
12. D. A. Rosenbaum, R. J. Meulenbrock, R. J. Vaughan, and C. Jansen, "Posture-based motion planning: applications to grasping," Psychol. Rev. **108**, 709–734 (2001).
13. U. Castiello, "The effects of abrupt onset of 2-D and 3-D distractors on prehension movements," Percept. Psychophys. **63**, 1014–1025 (2001).
14. J. Dean and M. Brüwer, "Control of human arm movements in two dimensions: paths and joint control in avoiding simple linear obstacles," Exp. Brain Res. **97**, 497–514 (1994).
15. L. A. Howard and S. P. Tipper, "Hand deviations away from visual cues: indirect evidence for inhibition," Exp. Brain Res. **113**, 144–152 (1997).
16. M. Mon-Williams, J. J. Tresilian, V. L. Coppard, and R. G. Carson, "The effect of obstacle position on reach-to-grasp movements," Exp. Brain Res. **137**, 497–501 (2001).
17. S. P. Tipper, L. A. Howard, and S. R. Jackson, "Selective reaching to grasp: evidence for distractor interference effects," Vision Cogn. **4**, 1–38 (1997).
18. D. A. Rosenbaum, R. J. Meulenbrock, R. J. Vaughan, and C. Jansen, "Coordination of reaching and grasping by capitalizing on obstacle avoidance and other constraints," Exp. Brain Res. **128**, 92–100 (1999).
19. A. F. C. Hamilton and D. M. Wolpert, "Controlling the statistics of action: obstacle avoidance," J. Neurophysiol. **87**, 2434–2440 (2002).
20. P. N. Sabes and M. I. Jordan, "Obstacle avoidance and a perturbation sensitivity model for motor planning," J. Neurosci. **17**, 7119–7128 (1997).
21. P. N. Sabes, M. I. Jordan, and D. M. Wolpert, "The role of inertial sensitivity in motor planning," J. Neurosci. **18**, 5948–5957 (1998).
22. C. M. Harris and D. M. Wolpert, "Signal-dependent noise determines motor planning," Nature **394**, 780–784 (1998).
23. E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," Nat. Neurosci. **5**, 1226–1235 (2002).
24. D. E. Meyer, R. A. Abrams, S. Kornblum, C. E. Wright, and J. E. Smith, "Optimality in human motor performance: ideal control of rapid aimed movements," Psychol. Rev. **95**, 340–370 (1988).
25. R. Plamondon and A. M. Alimi, "Speed/accuracy trade-offs in target-directed movements," Behav. Brain Sci. **20**, 279–349 (1997).
26. N. Smyrnis, I. Evdokimidis, T. S. Constantinidis, and G. Kastrinakis, "Speed-accuracy trade-offs in the performance of pointing movements in different directions in two-dimensional space," Exp. Brain Res. **134**, 21–31 (2000).
27. J. D. Connolly and M. A. Goodale, "The role of visual feedback of hand position in the control of manual prehension," Exp. Brain Res. **125**, 281–286 (1999).
28. M. Desmurget and S. Grafton, "Forward modeling allows feedback control for fast reaching movements," Trends Cogn. Sci. **4**, 423–431 (2000).
29. W. P. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, eds., *Numerical Recipes in C. The Art of Scientific Computing*, 2nd ed. (Cambridge U. Press, Cambridge, UK, 1992).
30. D. H. Brainard, "The psychophysical toolbox," Spatial Vision **10**, 433–436 (1997).
31. D. G. Pelli, "The videotoolbox software for visual psychophysics: transforming numbers into movies," Spatial Vision **10**, 437–442 (1997).
32. It is instructive to consider the extent to which visual feedback plays a role in our results. The movement times were 300–400 ms, which is barely enough time to allow for an influence of visual feedback during the movement. We ran one subject in a version of experiment 1 in which the visual stimulus disappeared as soon as the space bar was re-

leased. This manipulation eliminates feedback from the relative positions of the hand and the visible target during the movement but allows for feedback by using the view of the hand and apparatus. The results were unaffected, including the movement variability and the pattern of movement end points.

33. D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases* (Cambridge U. Press, Cambridge, UK, 1982).

34. In the 16 practice (warm-up) trials of this experiment, subjects pointed to each of the configurations twice in the penalty = 0 condition. Subjects were exposed to the eight novel configurations in the penalty = 500 condition for the first time during the experiment.