Testing Whether Humans Have an Accurate Model of Their Own Motor Uncertainty in a Speeded Reaching Task

Hang Zhang[1,2]*, Nathaniel D. Daw[1,2], Laurence T. Maloney[1,2]

Abstract

7 -300.81 -33943

Author Summary

? 50 20 *think*

and similar results found in the literature, it is tempting to assume that human movement planning is based on an accurate model of motor uncertainty.

One goal of the present study is to interrogate and ultimately challenge this assumption. One reason to do so is that previous evaluations of human performance are not sensitive to even gross errors in the representation of motor uncertainty. In Trommershäuser et al.'s [7] experiment, for example, human subjects' end points on the screen formed a bivariate Gaussian distribution centered at the aim point. Suppose a subject correctly estimates the variance of the end points but mistakenly assumes that the end points are distributed uniformly in a fixed circle around the aim point. This subject has a model of his own error distribution that is markedly different from her actual error distribution. The two distributions are illustrated in an inset to Figure 1B.

To evaluate the performance of such a hypothetical subject, we simulated the six reward conditions of Trommershäuser et al. [7] and plotted the results in Figure 1B. The differences between the optimal aim point (golden diamond) and the aim point of the hypothetical subject (black cross) is small, less than 1 mm on average and the average expected gain of the hypothetical subject was as high as 97.0% of the maximum expected gain. That is, although the hypothetical subject had an inaccurate model of her own error distribution, her performance would probably be

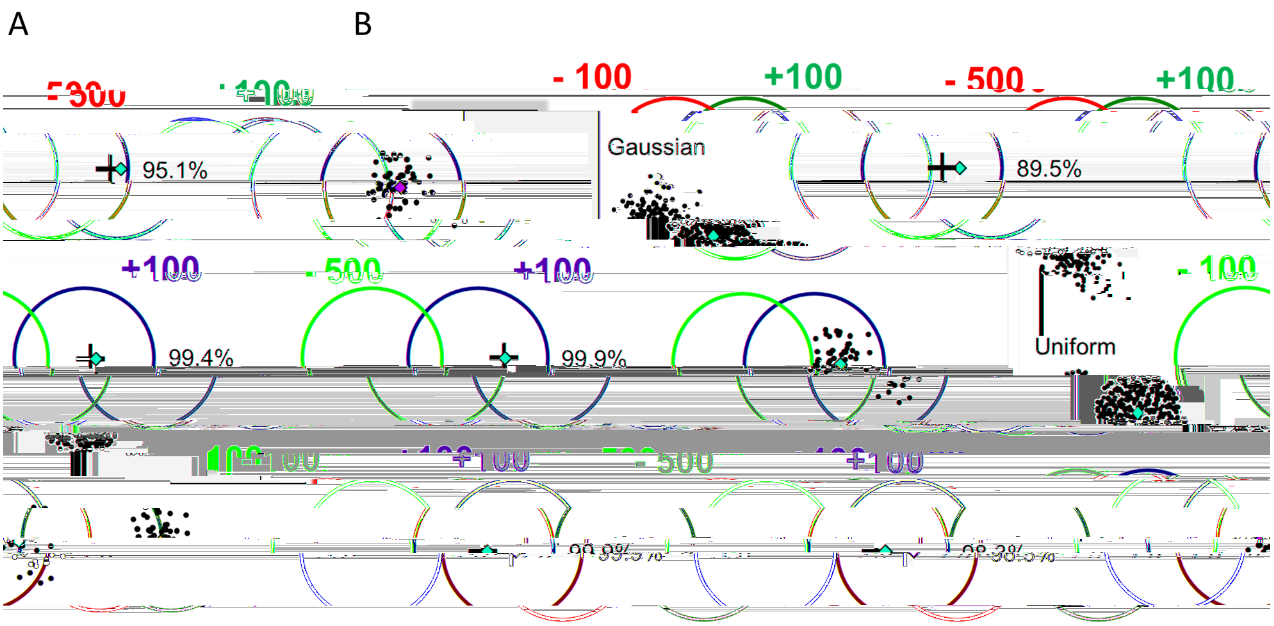

Figure 1. Performance of a hypothetical subject in Trommershäuser et al.'s [7] experiment. −100 −500, 1, 1.5, 2 8.97 +100. 0 4.05 0 4.05

:10.1371/ .1003080. 001

indistinguishable from optimal in Trommershäuser et al.'s [7] experiment, and in any of the studies we cited earlier.

We developed a simple motor choice task to more directly assess humans' internal models of their own motor error distributions. Human subjects were first trained to make speeded movements to radially-symmetric targets on a computer display. They were permitted only a short time to execute the movement and hit the screen. During the training, we estimated subjects' true motor error distributions $\phi(x,y)$. They were all well described as vertically elongated, bivariate Gaussian distributions.

In the second phase of the experiment, subjects did not attempt to hit targets. Instead they were given pairs of potential targets, one rectangle and one circle, of specific sizes. The task was to choose the target that was easier to hit (Figure 2A). Subjects knew that at the end of the experiment they would attempt to hit a small number of the targets they had chosen and they would be paid a cash reward for each success. The cash reward for either target was the same and it was therefore in their interest to choose the target in each pair that offered the higher probability of success.

If the targets are denoted T_1 and T_2 then the true probability of success in hitting the i^{th} target is

$$p_i \quad = \quad \int_{T_i} \phi(x,y)\,dx\,dy\;, \quad i=1,2. \tag{1}$$

The target is just the region of integration and the probability of success is just the proportion of the probability density function contained within the target. (We verified in training that subjects aimed at the centroid of the targets.)

But how is the *subject* to decide between targets? We consider the possibility that she has some internal estimate of the distribution of her own motor uncertainty, $\psi(x,y)$. In evaluating each target, she computes an estimate of probability based on this estimate,

$$p_i' \quad = \quad \int_{T_i} \psi(x,y)\,dx\,dy\;, \quad i=1,2, \tag{2}$$

and then chooses whichever target offers the higher probability.

This strategy would maximize expected gain in our task if the subject's estimate of her distribution were accurate: $\psi(x,y) = \phi(x,y)$. If it is not, then some of her choices will differ from the choice dictated by Eq. 1.

We illustrate the method by explaining how we test for isotropy (Figure 2B). Suppose the subject is given two rectangles, one horizontal and one vertical, of the same size and is asked to choose the one that is easier to hit. If the subject's internal model is an isotropic distribution, i.e. equal variance in the horizontal and vertical directions, the subject should be indifferent between the two rectangles. Alternatively, if the subject assumes a horizontally elongated distribution, i.e. a larger variance in the horizontal distribution, the subject would prefer the horizontal rectangle, and vice versa. (In practice, we never asked subjects to directly compare a horizontal rectangle and a vertical rectangle. Instead, we used a staircase method to determine the radii R_0 of the circle that the subject judged to be as "hittable" as any given rectangle and compared these equivalent radii.)

Ten (Experiment 1) or eight (Experiment 2) different rectangles, horizontal or vertical, were used and for each we measured its equivalent radius R_0, where the subject chose indifferently between the rectangle and the circle (Figure 2C). Based on a subject's choices for varying pairs of targets, we are able to test the variance and anisotropy of her distribution model.

In Trommershäuser et al.'s [7] task, subjects were facing "motor lotteries" with the probabilities of different outcomes determined by their own motor error. One concern is the possible effect of probability distortions on the interpretation of these studies. It is well-known that humans overweight small probability and underweight large probability in classical decision tasks [10], where they choose among economic lotteries. Wu, Delgado, and

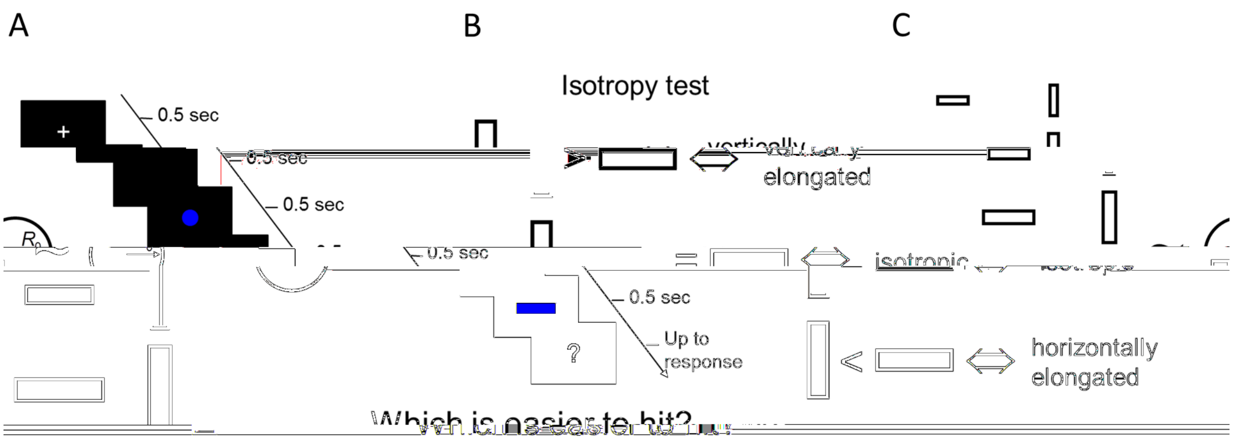

Figure 2. The choice task. A. ... ? 1 ... 2 ?". B. ... R_0. ... 2 ...

:10.1371/ ... .1003080. 002

Maloney [11] show that people have systematic probability distortions with motor lotteries as well, although in a reverse pattern: they underestimated small probabilities and overestimated large probabilities. Choice between targets in our task depends only on ordering of the estimates of the probabilities of hitting them – since the reward associated with success never varies – and is thus insensitive to any distortion of probability. In particular, if $w(p')$ is any strictly increasing function of probability p' which the subject applies to the probabilities computed from Eq. 2, then $w(p'_1)=w(p'_2)$ precisely when $p'_1=p'_2$.

Results

We report the results of two experiments. For simplicity, we focus on Experiment 1 and use Experiment 2 to address concerns raised in Experiment 1. We first describe the true motor error distributions measured in the training task. We then fit subjects' responses in the probability choice task to a probabilistic model with two free parameters and compared subjects' models to those of their true motor error distribution. We report large, systematic deviations. Particular patterns in subjects' model failures are identified. Unless otherwise stated, the significance level used was .05.

, I 1

True motor error distribution. The true error distribution depends on the trajectory of the movement, which we controlled by ensuring that participants started all reaches from a common starting point (the space bar of our computer keyboard). Additionally, the subject has some control over her motor error distribution $\phi(x,y)$. Normally, she can move more or less quickly, altering $\phi(x,y)$, and previous research demonstrates that humans do trade speed for accuracy, information or reward [2–4]. By imposing a time limit on the movement (from release of the space bar to touch of the screen) we effectively eliminated this freedom.

Only movements completed within the time limit were included into analysis. The end points for a typical subject are shown in Figure 3A.

We first examined Q-Q plots of all subjects' end points; as in past work, the Q-Q plots were close to linear, indicating that subject's motor errors were close to bivariate Gaussian. In Figure 3B we show the Q-Q plots for one typical subject.

For each subject, we fitted the x and y coordinates of the endpoints to a bivariate Gaussian distribution. For simplicity, we treated the x and y errors as independent (uncorrelated) and estimated σ_x and σ_y by computing the standard deviation separately for the horizontal and vertical direction. However, for 10 of the 18 subjects, the errors in the x and y directions were significantly correlated (ρ ranged from -0.44 to 0.33). We verified that the slight "tilt" introduced by correlation (e.g. Figure 3A) had negligible effect on any of our further tests: Taking into account the correlation would have no influence on the probability of hitting circles and would change the probability of hitting any of the rectangles we tested by no more than 3%.

Figure 3C shows the relationship between σ_x and σ_y. We tested for equality of variance separately for each subject (one-tailed F test). All subjects had a significantly larger σ_y than σ_x: the distribution was vertically elongated. The median σ_y-to-σ_x ratio across subjects was 1.44.

In summary, subjects' estimated motor error distributions were vertically elongated bivariate Gaussians. If we define the *variance parameter* $\sigma_0=\sqrt{\sigma_x\sigma_y}$ and the *anisotropy parameter* $\alpha_0=\sigma_y/\sigma_x$, the motor error distribution can be written as:

$$\phi(x,y)=\frac{1}{2\pi\sigma_0^2}\exp\left(-\frac{x^2}{2\sigma_0^2/\alpha_0}-\frac{y^2}{2\sigma_0^2\alpha_0}\right) \quad (3)$$

Subjects' model: Gaussian vs. area-matching. To allow for concrete parametric comparisons, we estimate the subjects' models of their own motor error distributions, $\psi(x,y)$, assuming they have the same Gaussian form as $\phi(x,y)$ but with possibly different variance and anisotropy parameters:

$$\psi(x,y)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2/\alpha}-\frac{y^2}{2\sigma^2\alpha}\right) \quad (4)$$

(We consider the possibility of other distributional forms in the discussion.) Based on their choices, we estimated each subject's σ and α (see Methods). Figure 4A shows the estimates of σ and α, relative to σ_0 and α_0.

We considered the possibility that some subjects were choosing the target of larger area rather than the target of larger probability of a hit. If a subject used this area-matching strategy, his estimated σ would approach infinity. Indeed, the estimated σ of a considerable number of subjects were far greater than σ_0, up to 889 times of σ_0.

For each subject, we tested the Gaussian model (Eq. 4) against an area-matching model (see Methods). The area-matching model could be treated as a special case of the Gaussian model with $\sigma\rightarrow\infty$ and $\alpha=1$. According to nested hypothesis tests [12], 10 out of 18 subjects were better fit by the Gaussian model and the remaining 8 subjects were better fit by the area-matching model. For convenience, we call the former the Gaussian type, the latter the area-matching type.

Subjects' model: variance and anisotropy. We explored the variance and anisotropy of the subjects of the Gaussian type (Figure 4A, in green; we did not further examine the parameter estimates for subjects of the area-matching type). If a subject's model were the same as her true motor error distribution, her data point in Figure 4A would fall on the coordinate (1, 1). According to the 95% confidence intervals of $(\sigma/\sigma_0,\alpha/\alpha_0)$, all the subjects' models deviated from their true motor error distributions.

For a bivariate Gaussian distribution, the central regions have higher probability density than peripheral regions. Because circles are more concentrated than rectangles, a circle that is as equally "hittable" as a specific rectangle should be smaller than the rectangle in area. Intuitively, the larger the variance parameter σ, the more dispersed the assumed distribution, the larger the equivalent radius for a specific rectangle.

For most of the subjects of the Gaussian type, the estimated internal variance was close to their true variance (Figure 4A). For 4 out of these 10 subjects, σ/σ_0 was not significantly different from one. Four subjects' σ/σ_0 were significantly less than one and two subjects, significantly greater than one.

The anisotropy parameter α determines the perceived relative "hittability" of horizontal and vertical rectangles, as we illustrated in the Introduction. If two rectangles have the same size, the larger the α, the larger the equivalent radius of the vertical rectangle relative to that of the horizontal rectangle.

All the subjects of the Gaussian type underestimated the vertical anisotropy of their true distribution, with all their α/α_0 significantly less than one. We plotted α against α_0 (Figure 4B) and further examined whether subjects' α was sensitive to their true anisotropy α_0. Subjects' α_0 varied from 1.14 to 1.64. There was no significant correlation between α and α_0, Pearson's $r=.34$, $p=.34$.

Instead, α was always close to one. For 8 out of the 10 subjects, the α was indistinguishable from one. That is, most of the subjects of the Gaussian type incorrectly treated their error distribution as isotropic.

To summarize, there were two patterned biases in subjects' models in the probability choice task: First, approximately half of the subjects failed to take their own motor error distributions into account and evidently based their choices on the areas of the targets instead. Second, among the subjects who correctly assumed a Gaussian model, 4/5 of them incorrectly assumed the vertically-elongated distribution to be isotropic.

Results of the area choice task. As a control for the probability choice task, in a subsequent area choice task, subjects were asked to choose which target was larger in area. We investigated whether the Gaussian and area-matching subjects also differed in their judgment of area.

In the area choice task, the equivalent radii of subjects of the area-matching type were close to the true area equivalent radii (i.e. area-matching) while subjects of the Gaussian type had smaller equivalent radii than the true (Figure 5B). This separation resembled that in the probability choice task (Figure 5A).

For probability choice, the reason equivalent radii are expected to be smaller than true radii is that the probability of hitting a circle is larger than hitting a rectangle equal in area: a circle fits more compactly near the center of an isotropic Gaussian error distribution. This effect reflects the implicit motor model because the radius difference is most dramatic for sharper Gaussians (smaller σ/σ_0) and vanishes in the limit of the infinite variance Gaussian (i.e., area matching). Accordingly, that their equivalent radii are smaller even in the area choice task suggest that subjects of the Gaussian type compensated for the probabilities of hitting targets even when judging area, although not as much as they did in the real probability choice task: their median σ/σ_0 was 0.97 in the probability choice task and 2.29 in the area choice task.

Indeed, we found that 8 out of 10 Gaussian subjects' area choices were better fitted by the Gaussian model than by the (now appropriate) area-matching model. In contrast, only 2 out of 8 area-matching subjects were better accounted by the Gaussian model. Subjects of the Gaussian type had a significantly larger proportion to incorrectly assume Gaussian in the area choice task, according to a Fisher's exact test.

We verified that subjects were not just confusing the probability and area choices and they did have different equivalent radii (R_0) in the two tasks. For each subject, we

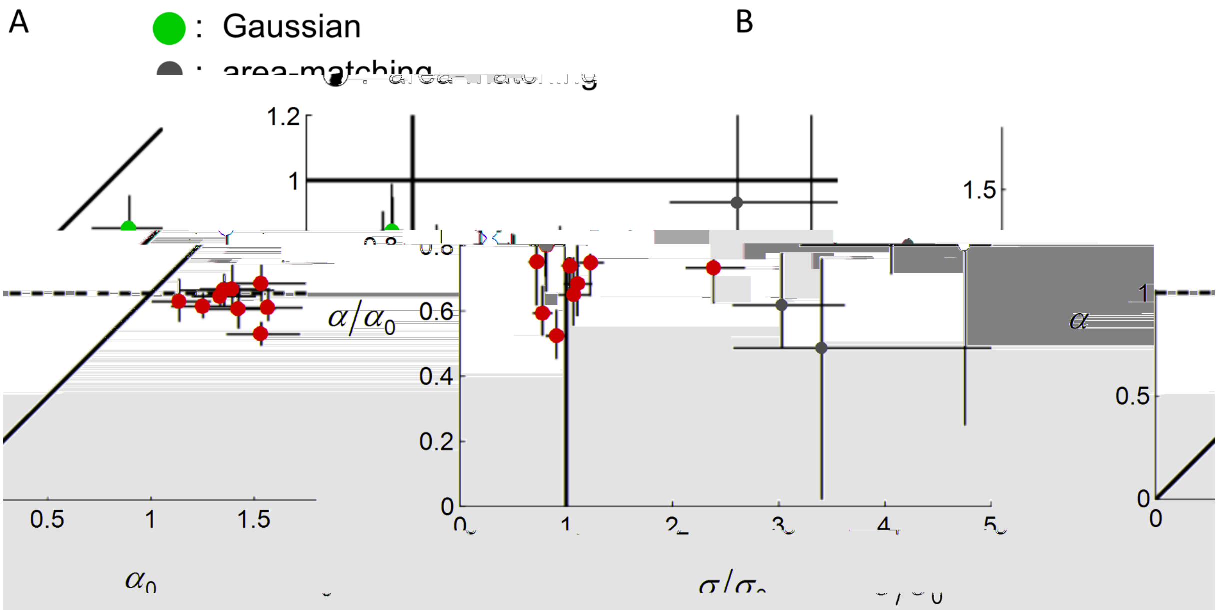

Figure 4. Subjects' models in probability choices in Experiment 1. 95% . σ_0 α_0 $\phi(x,y)$ (. 3). σ α $\psi(x,y)$ (. 4). A. α/α_0 σ/σ_0. Al 18 , 10 ( ) (σ/σ_0 1) ($\alpha/\alpha_0 < 1$). 8 ( ) σ/σ_0 (13, 19, 57, 889) α α_0 α 1 α_0

:10.1371/ .1003080. 004

Gaussian type. The remaining 3 subjects of the area-matching type had the probability R_0's significantly larger.

2

In Experiment 1, subjects' models of their own motor error distributions had patterned deviations from the true distributions. In Experiment 2, we tested whether these deviations could be eliminated or reduced by two manipulations.

First, we tested the effect of having more reaching experience with a single target, which might be expected to improve subjects' motor models. In the training phase of Experiment 1, subjects made 300 speeded reaches to a circular target of fixed size. We

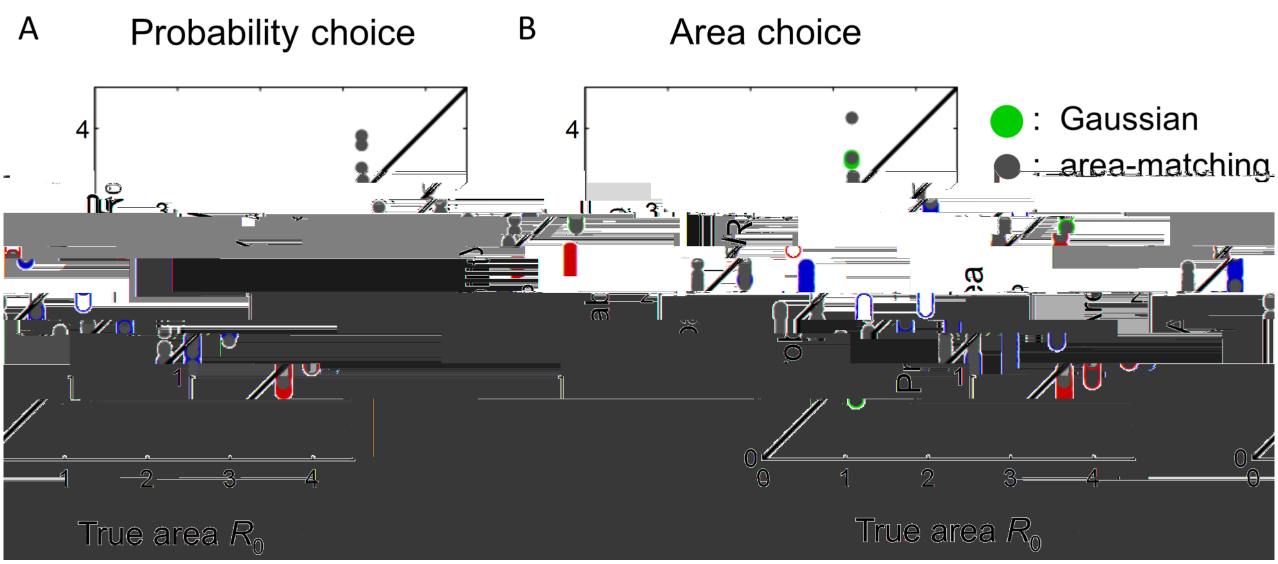

Figure 5. Probability vs. area choices in Experiment 1. (A) (B)

:10.1371/ .1003080. 005

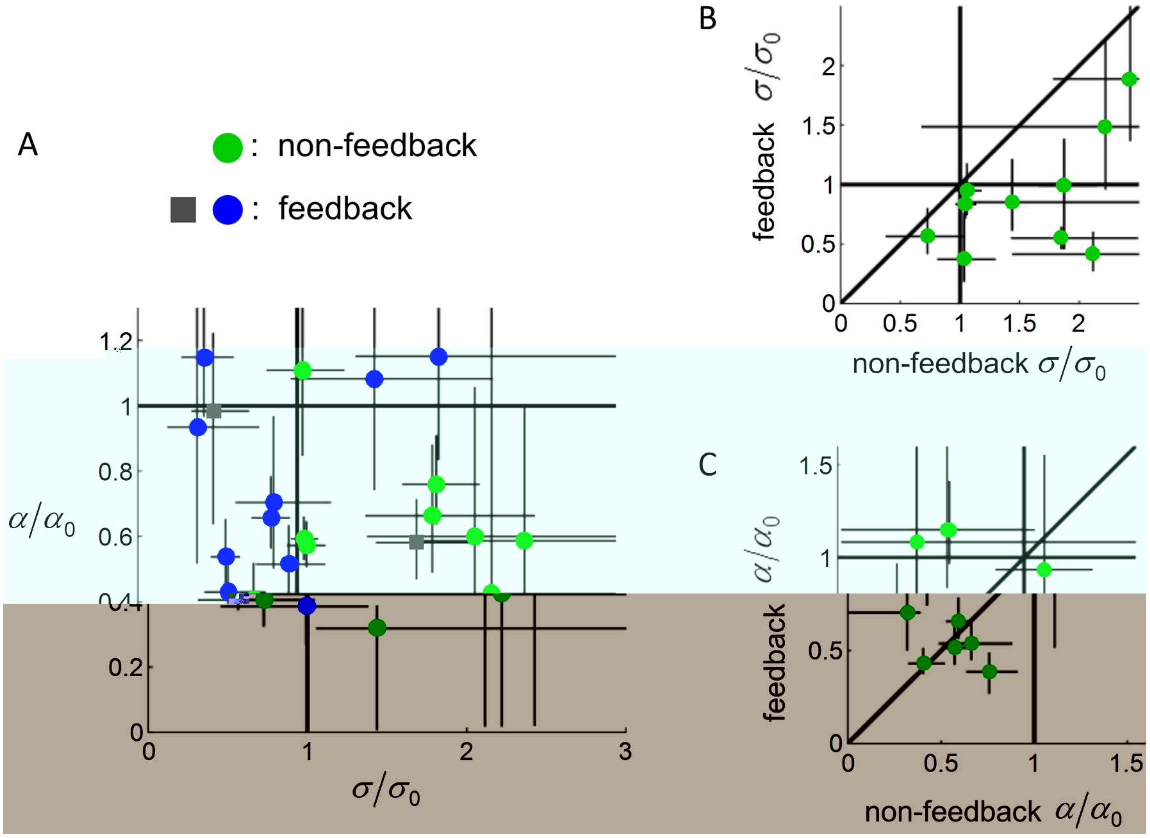

Figure 6. Subjects' model in Experiment 2. A. α/α_0 … σ/σ_0. … (…), … α/α_0 … 15 … 1210, …). B. … B. σ/σ_0 … σ/σ_0 … α/α_0 … α/α_0 … 95% …

:10.1371/ … .1003080.006

doubled the training trials to 600 in Experiment 2. As in Experiment 1, all subjects' true distributions were vertically elongated, bivariate Gaussian distributions.

Even with this more extensive training, among 12 new subjects, there were 2 subjects who were better fit by the area-matching model. Although this proportion is numerically smaller than that in the first experiment, the difference was statistically insignificant, according to a Fisher's exact test.

For the 10 subjects of the Gaussian type (green circles in Figure 6A), 5 subjects' σ/σ_0 were indistinguishable from one and 5 subjects' σ/σ_0 were significantly greater than one.

As in Experiment 1, most of these subjects, though not all, underestimated the vertical anisotropy of their distributions. Six out of the 10 close-to-one subjects' α/α_0 were significantly less than one (Figure 6A). Four subjects' α/α_0 were indistinguishable from one, but among them, three had α/α_0 between 0.42 and 0.60 – the inability to reject the null hypothesis was probably just because of the inaccuracy of the measurement (see Figure S1 for an illustration).

The second manipulation we added in Experiment 2 was to include a final experimental phase in which subjects would choose which of two targets they would prefer to hit and, immediately after choosing the target, subjects would actually make a speeded reach to the target and watched its consequence. This task gave subjects an opportunity to observe the probability of hit of targets of different shapes and sizes. It came after the probability choice phase (which was conducted, as before, with no feedback), so it would have no influence on the results of the latter that we reported above.

In the feedback phase, none of the 12 subjects were better fit by the area-matching model. All the subjects' σ/σ_0 were less than those in the non-feedback task (Figure 6AB). Six of the 12 subjects now significantly underestimated their variance.

As to α/α_0 (Figure 6C), 4 out of the 6 subjects who significantly underestimated the vertical anisotropy in the non-feedback task had no improvement (those whose error bars cross the identity line) and one subject performed even worse (those whose error bars are under the identity line).

To summarize, receiving feedback on a variety of targets appears to correct the area-matching strategy but does not improve underestimation of variance or anisotropy.

Discussion

We reported two experiments testing whether humans have accurate internal models of their own motor error distributions. In the first part of the experiments, subjects executed the reaching movements hundreds of times, allowing us to measure the error distribution of their end points. Though differing in variance, the distributions were all well characterized as bivariate Gaussian, all elongated in the vertical direction. Subjects were moving from a keyboard below the screen to the screen. A larger variance along the direction of movement (upward) is often reported [13–15].

Next, subjects were asked to repeatedly choose between two targets, selecting the one that appeared easier to hit. Based on their choices we tested two aspects of their internal model of their own distribution, variance and anisotropy.

The two experiments led to converging results: More than half of the subjects had accurate or almost accurate estimation of their own variance (no more than twice and no less than half of the true variance), while the rest failed badly, markedly overestimating their own variances or not taking into account their variance at all. Almost all the subjects failed to have an accurate estimate of how their distribution was shaped, incorrectly assuming a more isotropic distribution.

These failures are unexpected. Previous studies have shown close-to-optimal compensation for motor uncertainty [2–8]. In most of the studies [e.g. 7], subjects' decision under motor uncertainty was indicated implicitly by their movement. But close-to-optimal compensation was also observed in tasks resembling ours where explicit choices between two alternative options were required [16]. Patterned failure in compensating for motor uncertainty has seldom been reported.

As an exception, Hudson, Tassinari, & Landy [17] added anisotropic noise to the visual feedback of subjects' movements and found that people ignored the induced anisotropy. Our results in testing isotropy are consistent with theirs. But in their case, the artificial visual feedback conflicted with subjects' sensorimotor feedback. It is possible that subjects were just giving a higher trust to their own sensorimotor feedback. Our results are free of such possibilities: people do not correctly compensate for anisotropy even when the anisotropy emerges naturally.

The finding that a considerable proportion of subjects did not base their choices of movements on a model of their own distribution is unexpected and striking. However, it is not necessarily in conflict with previous studies where human performance is found to be close to optimal [7]. In previous studies, subjects performed real movements and received feedback in the test task as in the feedback task of our Experiment 2, where no subjects followed the area matching strategy.

In specific task situations, people can compensate for a missing or inaccurate model of their motor error distribution by using intuitive strategies. This is demonstrated in an anomaly reported by Wu et al. [18]. They used the same task as Trommershäuser et al. [7] with a different reward landscape and found sub-optimal human performances. In Trommershäuser et al. [7], the optimal aim point fell on the symmetrical axis and within the rewarding circle, both of which were highly intuitive. In contrast, Wu et al. [18] used an asymmetrical reward landscape that consisted of one rewarding circle and two penalty circles. The optimal aim point fell within one of the penalty circles. Subjects' failure in this counter-intuitive situation suggests that apparently optimal performances may rely on simple intuitive strategies.

An underestimation of variance, observed for half of the subjects even with feedback (Experiment 2), is probably not as costly to the same extent as an overestimation in the sorts of tasks considered by Trommershäuser and colleagues. Subjects in Trommershäuser et al.'s [7] task also received feedback after every trial and potentially this led to underestimation of variance in that experiment. We considered whether a considerable underestimation of their own motor variance could be compatible with humans' close-to-optimal performances. As we stated in the Results, in Trommershäuser et al.'s [7] task, if subjects overestimated their variance to up to 4 times the true variance ($\sigma/\sigma_0 = 2$), their expected gain would be only 74% of the maximum expected gain. To our surprise, however, if subjects underestimated their variance to 1/4 of the true variance, ($\sigma/\sigma_0 = 1/2$), their expected gain would be as high as 96% of the maximum expected gain. That is, at least in this situation, an underestimation of variance incurred little penalty.

Subjects' failures in our experiment could not be attributed to a mere lack of experience. Before the two choice tasks, subjects repeated the reaching movements for over 300 or 600 times and the position of the endpoint was provided after each reach. Moreover, goal-directed reaching is arguably among the most practiced motor tasks of everyday life. Therefore, it is a mystery that people are not able to model their own motor error correctly and exhibit considerable and patterned deviations. Whence come their incorrect models? When will incorrect models be abandoned and be replaced by the correct ones?

For example, why did subjects assume an isotropic distribution? We conjecture that they were trying to use a model that is as simple as possible to fit their observations. When observations are few even a real isotropic distribution may be better fit by an anisotropic model. Thus, by adopting a simpler model, subjects could avoid over-fitting their observations. The problem is: why should people stick to an incorrect model even after hundreds of observations?

Another intriguing fact is the double dissociation between subjects of the Gaussian type and subjects of the area-matching type. For the area-matching subjects, area-matching seems to be their substitute strategy for the probability choice task; while for the Gaussian subjects, "probability-matching" seems to be their substitute strategy for the area choice task. If we consider the area of a target to be the integration of a unit probability density across the target region, area choices are comparable to probability choices. Is there any common process involved?

The choice task we designed in the present study is a powerful tool for determining what people "know" about their motor error (or more precisely, what model of motor error is consistent with their choice performance). It only asks for an ordering of probabilities. It does not depend on a utility function, since the two targets they are choosing from are associated with the same amount of reward. It is not influenced by how probability is non-linearly distorted [10,19], so long as the distortion function retains the order of the probability scale.

Our task differs from most other motor decision tasks [e.g. 2,7,9] in two respects that might in principle produce different results: the choices are binary rather than continuous, and concern hypothetical future rather than immediately actualized movements. For example, people may not have full access to their motor uncertainty in the absence of real movement planning or execution. The existing evidence, however, does not suggest human choices in a binary, hypothetical motor task would necessarily differ from those in continuous, actualized movements. First, close-to-optimal performances were found in previous studies on binary, hypothetical motor decisions [16]. Second, the neural circuits activated by real and imagined movements are highly similar [20,21].

In the present study, we estimated subjects' behavior using a Gaussian distributional assumption to allow direct quantitative comparison with the ideal observer model in terms of that model's parameters. Of course, it is possible that subjects assumed a different distributional form subjectively, or even chose based on some heuristic that does not directly correspond to the decision theoretic model for any distribution. Although both of these possibilities are interesting hypotheses for the source of the sub-optimality we reveal, even if true they would not invalidate the results of the present analysis in using the Gaussian fits descriptively to characterize the existence and nature of deviation from the ideal observer. As pointed out by Geisler [22], it is valuable to compare actual to ideal even when people are not ideal.

What distinguishes our study from previous studies is an exploration of the most likely model implicit in each individual's performances. We broke down the ideal observer into multiple dimensions (variance and anisotropy) and assessed human observers on these dimensions. The multi-dimensional tests accommodate the possibility that a specific individual may deviate from the ideal observer on some dimensions but not others, which a one-dimensional test would not afford. The deviation on each dimension is separable in subjects' choices. Our task is thus sensitive to the each particular individual's possible deviations from ideal and provides alternative models to ideal.

In a recent article [23] we found people do not have an accurate model of their own visual uncertainty. Subjects chose between visual discrimination tasks that could differ in location (retinal eccentricity) and contrast. By examining subjects' choices we could test what they implicitly assumed about their own retinal sensitivity in the periphery. We found that all but one subject was not even consistent in their choices: the pattern of choices violated transitivity of preference, i.e. in some cases they preferred lottery A over lottery B and lottery B over lottery C but, finally, lottery C over lottery A.

Had we simply compared subjects' performances to optimal in Zhang et al [23] and the present paper, we would only have concluded that subjects' performance was less than ideal, overlooking the striking patterns of failure and individual differences that we instead found.

Methods

The experiment had been approved by the University Committee on Activities Involving Human Subjects (UCAIHS) of New York University and informed consent was given by the observer prior to the experiment.

preferred. They were rewarded for hits on these trials just as in the training task. The eight targets were randomly selected from the targets that they had judged to be easier to hit.

Area choice. The area choice task was a control to the probability choice task, where the same stimuli were used but now the task was to choose the alternative that was larger in area.

Subjects were rewarded for correct responses. Subjects were instructed that eight trials would be selected at random after they completed the 1000 trials. For each of the trials, subjects would win $1 if their choice had been correct.

l 2

Subjects. Twelve new naïve subjects, six female and six male, 11 right-handed and one left-handed, participated.

Apparatus and stimuli. Same as Experiment 1, except that no chinrest was used.

Procedure and design. Subjects completed three tasks in one session in the following order: *training*, *probability choice*, and *probability choice with feedback*. The first two tasks were the same as those of Experiment 1. The task of probability choice with feedback was similar to a combination of the tasks of probability choice and training. On each trial, subjects first chose between two sequentially displayed targets the one that was easier to hit. After that, they initiated a pointing trial by placing their index finger on the space bar and would try to hit the target they chose within the time limit. Feedback of the endpoint was given as in the training task. The bonus rule was the same as that of the training task.

In the two tasks of probability choice, same as Experiment 1, rectangles' sizes were tuned to each subject's motor variance and the radius of the circle was adjusted by adaptive staircase procedures. The design differed from Experiment 1 only in the sizes of rectangles and settings of staircase procedures. There were 8 (2 orientations by 4 sizes) different rectangles, whose short-side lengths were: σ_0, $1.40\sigma_0$, $1.96\sigma_0$, $2.74\sigma_0$. For each of the 8 rectangles, there was one 1-up/1-down staircase that terminated after 60 trials. (The motivation for us to adopt the one 1-up/1-down staircase for each rectangle rather than the one 1-up/2-down and one 2-up/1-down staircases in Experiment 1 was to use fewer trials to estimate each equivalent radius. This change would not introduce any known biases into the estimation of equivalent radius, given the data fitting procedures described below.) The step sizes were 0.15, 0.1, 0.08, 0.06 in log units, respectively for the first, second, third and the remaining reversals. Interleaved, the 8 staircases $\times 60 = 480$ trials were run in blocks of 60 trials, preceded by 8 warm-up trials.

Equivalent radius. For a specific rectangle in a choice task, we defined the *equivalent radius* (R_0) as the radius of the circle such that the subject was indifferent between the rectangle and the circle in her choice. The method of eliciting the equivalent radius was as follows.

For each subject and each specific rectangle, we assumed that the probability of choosing the circle was a Quick-Weibull psychometric function [26,27] of the radius of the circle R:

$$F(R) = 1 - \exp\left(-\left(\frac{R}{\lambda}\right)^{\gamma}\right) \quad (5)$$

where λ is a position parameter and γ is a steepness parameter. Assuming that different rectangles were associated with different λ but a common γ, we estimated λ and γ for each rectangle by fitting the responses of the staircase trials to Eq. 5 using maximum likelihood estimates.

By the definition of equivalent radius, 1 6 0 (l d 7 (F r R u n) - 2 5575(7a8(s if ze ee sd)b this test statistic is asymptotically distributed as a χ^2 random variable with degrees of freedom equal to the difference in number of parameters in the two models under comparison. Accordingly we compared Δ to the 95$^{\text{th}}$ percentile of a χ_2^2 distribution.

Confidence intervals. We computed the 95% confidence intervals of σ/σ_0, α/α_0, α, α_0 using a bootstrap method [28]. For each subject, we ran a virtual experiment for 1000 times and estimated the above measures on each run. In the training task, endpoint positions were resampled from the non-time-out trials. In the probability choice task, the responses of each staircase trial was generated by parametric resampling [29] from the psychometric