DECISION MAKING AND MOVEMENT PLANNING UNDER RISK

Julia Trommershäuser[1,2], Laurence T. Maloney[2], Michael S. Landy[2]

1: Department of Psychology, Giessen University, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany;
2: Department of Psychology and Center for Neural Science, New York University; 6 Washington Place, New York, NY 10003, USA

Motor variability makes movement outcome uncertain. In past work (Trommershäuser et al., 2003, *JOSA A*, *20*, 1419), subjects rapidly touched a target while trying to avoid a nearby penalty region. As we varied reward, penalty, and target/penalty locations, subjects modified their movement plans to maximize expected gain (EG). They effectively considered each possible outcome (penalty, reward, both, neither) paired with the probability that the planned movement would lead to that outcome. Such a list of paired outcomes and probabilities constitutes a *lottery*. We found that subjects chose the lottery with maximum EG. In contrast, humans typically fail to maximize EG when choosing lotteries presented in pencil-and-paper form (Kahneman & Tversky (Eds.), *Choices, Values & Frames*, 2000, Cambridge Univ. Press). In our task, probabilities are due to the subject's intrinsic motor variability, while in pencil-and-paper tasks they are chosen by the experimenter and conveyed symbolically.

In our previous experiments, subjects adjusted their movements to maximize EG in response to changes in *value* (penalty values of 0, -100, -500). In another experiment, we manipulated task-relevant movement variability, and thus the probability (*uncertainty*) of incurring rewards and penalties. Subjects compensated for changes in their uncertainty to maximize EG (Trommershäuser, Gepshtein, Maloney, Landy & Banks, under review).

Here, three experiments were performed to explore the link between human movement planning under risk and human decision making. Subjects attempted to rapidly touch a target region displayed on a computer screen, while trying to avoid a nearby penalty region. Each target hit yielded a monetary reward; each penalty hit yielded a monetary penalty. Responses later than 700 ms after a start signal were penalized. Subjects performed an initial practice session to learn these time constraints (with different configurations than those used in the experiment). Once subjects met these time constraints, the rapid execution of pointing movements resulted in substantial movement variability. (See Trommershäuser et al., 2003, for details apparatus and procedure, and computations of predictions based on EG maximization.)

In the first experiment, we modified our task so that, in some conditions, the reward and/or penalty regions were stochastic: when the subject hit a stochastic region, s/he would receive the reward or penalty with probability 0.5. We compared performance for all four combinations of certain or stochastic rewards and penalties. Each condition was run in a separate session and the first and last sessions were always fully certain. The subject was explicitly told the probabilities before each session. Six naive subjects completed the experiment. Five out of six subjects maximized EG in the fully certain conditions (in both the first and the last session). However, five of the six subjects were markedly sub-optimal in one or more of the stochastic conditions. In the majority of cases, subjects shifted their movement end points in response to manipulations in explicit probabilities in the sub-optimal direction, i.e. closer towards the penalty when the chance of scoring a reward dropped to 50%, or further away from the penalty when the chance of scoring a penalty dropped to 50%. In the condition in which both the penalty and reward were stochastic, the optimal strategy is identical to that in the certainty condition. Yet, four out of six subjects changed their mean movement end points significantly in this condition. However, all subjects returned to their initial optimal strategy in session five when rewards and penalties were scored with certainty. In summary, associating extrinsic probabilities with rewards and penalties disrupted optimal performance.

In the second experiment, we varied the time between stimulus display and the start signal for the pointing movement. The stimulus configuration was displayed simultaneously followed by a 400 Hz tone (0 ms, 400 ms or 1000 ms after stimulus display) indicating the start signal for the movement. Six subjects completed the experiment. Subjects' movement end points and performance remained unaffected by variations in stimulus presentation time. Thus, human movement strategies remained stable (and did not differ significantly from optimal) across variable presentation times of the available rewards and penalties. Having more time to think about the decision did not disrupt optimality.

In a third experiment, there were two configurations, each containing a reward and a penalty region. In some trials, the two configurations differed with respect to penalty value and spatial arrangement of the reward and penalty regions. Thus, the two configurations, by themselves, had different values of maximum EG. Five subjects were instructed to rapidly point at one of the two simultaneously displayed stimulus configurations. 400 ms after stimulus display, a tone indicated that the subject should initiate the movement. Subjects' movement end points that ended in one sub-configuration did not differ from the distribution of movement end points in trials with only a single configuration (control experiment). In trials with two configurations with different maximum expected gain, four of five subjects pointed more frequently at the configuration with higher maximum expected gain (preferences of 81% to 55%). Overall performance was suboptimal for three out of five (right-handed) subjects due to a preference for the stimulus configuration presented in the right half of the screen (81% to 72% across the overall balanced design). Thus, human movement strategies remained stable (and optimal) for these selection movements that indicated a choice among multiple configurations. But, the choice among configurations with different expected gain was sub-optimal due to constraints of the motor system.

While we have argued here that movement under risk is formally equivalent to decision making under risk and uncertainty, many aspects concerning presentation, execution, completion and repetition of the task, the feedback about possible errors, as well as the implementation of the pay-off rule differ between the two approaches. Thus, cognitive, perceptual and motor constraints come into play when deciding on an action and subsequently planning and executing a goal-directed movement under risk.