UNDERGRADUATE THESIS:

STRIATAL ACTION-COMPETITION FRAMEWORK ACCOUNTS FOR DECISIONS IN PERCEPTUALLY UNCERTAIN MOTOR GAMBLES

RORY FLEMMING

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

DEPARTMENT OF NEUROSCIENCE

Abstract

How does the brain combine sensory and economic information when making value-based sensorimotor decisions? To investigate this problem, we designed a visually-guided sensorimotor gambling task to force human participants to consider both the economic value of targets as well as sensorimotor uncertainty in their selections to maximize points accrued. Over a series of three experiments, participants (N=60) selected risky spatial targets more frequently over safer targets that delivered a constant but small reward, if the reward magnitude for the risky target was high relative to a possible penalty magnitude, if the reward probability was high for the risky target, and if the spatial variance of the safe target was high (i.e., high selection uncertainty). These target parameters interacted to drive participants' decisions. The Believer-Skeptic model of basal-ganglia dependent decision-making predicts that economic value and sensory uncertainty should impact the drift rate of the decision process, instead of the decision threshold. This was confirmed using hierarchical drift-diffusion models. This suggests that sensory and economic information about the perceptual motor gambles are combined through the dynamic competition between the direct and indirect pathways associated with each gamble.

Keywords: risk, perceptual uncertainty, motor gamble, decision making, drift-diffusion model, believer-skeptic framework

INTRODUCTION

In a heated game of softball, an outfielder struggles to get a handle on the ball after a double is fired off. Prior to this there were runners on first and second. The outfielder could throw to home to out the lead runner and prevent a loss at the risk of letting both runners through, or throw to the pitcher to end the play as soon as possible. The outfielder only has enough time to fixate on one target before launching the ball. Where should the outfielder look and throw? Many value-based decisions, such as gambling on a risky spatial target with a potentially high reward, should be contingent on the saliency of sensory signals. These sensory signals should interact with feedback-driven learning from past choices; however, it remains unclear how sensory evidence may impact these types of risky decisions.

Economic decision theories attempt to answer how humans make decisions under risk by focusing on the expected value of decisions. Gambles may pit two or more actions with uncertain outcomes against one another. The classical cumulative prospect theory describes human valuation of economic gambles using nonlinear weighted probability and utility functions to calculate the gamble with the highest utility (Tversky & Kahneman, 1992). The theory predicts that people choose gambles with the highest utility. Using humans as the basis for the probability and utility functions generates the predictions that people tend to be more risk aversive, since losses are weighted more heavily than rewards. However, if the magnitude of the potential reward is significantly larger than the potential loss, people are more willing to take risks. These theories help to describe how choices are valued, but do not propose biological mechanisms by which decisions are executed.

One recently proposed biological hypothesis for understanding how actions are selected is the Believer-Skeptic framework. This framework proposes that uncertainty in actions is

encoded in the competition between the direct and indirect pathways of the basal ganglia (Dunovan & Verstynen, 2016). The direct pathway of the basal ganglia is responsible for the initiation of actions while the indirect pathway suppresses alternative actions (Mink 1996). In the Believer-Skeptic Framework, the direct (Believer) and indirect (Skeptic) pathways are modeled as mutually inhibiting, recurrently exciting accumulators. Evidence for a decision is input to the system. As more evidence is accumulated for a decision, the direct pathway activity ramps up to enforce the supported choice, while the indirect pathway acting on other actions also ramps up to suppress these alternatives. Imagine being presented a choice between a risky decision and safe decision. As your brain considers the benefits of making each choice, the direct pathways and indirect pathways associated with these two choices will be activated. When the choices are similar in value to you, it is a more challenging decision, and there is more uncertainty. This results in more narrow competition within the basal ganglia and a slower reaction time since you will likely have to consider more evidence before making a decision. However, if the decision is easy, the direct excitation for the preferred decision and indirect inhibition of alternatives is so strong, the decision can be made more quickly. The Believer-Skeptic framework gives biological constraints to quantitative models of decision making, allowing for better understanding of the relationship between implementational mechanisms and decision algorithms.

In order to effectively apply and interpret the Believer-Skeptic framework, it is necessary to identify the inputs to a given decision. Two factors have been known to impact spatially guided sensorimotor decisions: perceptual uncertainty and expected gain. Noise within and external to the nervous system requires that the brain estimate and account for such noise as it affects goal outcomes (Faisal, Selena & Wolpert, 2008). Perceptual discrimination has been studied from the level of individual neurons to organism behavior, including non-human

primates and humans (Britten et al., 1996). Evidence suggests that people learn to make movements that minimize the error in outcomes by combining current and prior information (Körding & Wolpert, 2004), and when relevant, by accounting for how to maximize gains acquired from the task at hand (Trommershäuser, Maloney & Landy, 2003). Visual uncertainty is known to directly increase associated motor variability (Körding & Wolpert, 2004). Therefore, visually-guided sensorimotor decisions depend on the perceptual and motor uncertainty as they affect the expected gains associated with their actions.

One way of investigating the effects of sensorimotor noise and expected gains on human selection behaviors is through the implementation of target estimation tasks. For example, Trommershäuser, Maloney & Landy (2003) had subjects rapidly point at partially overlapping target circles, and demonstrated the people plan movements to choose an average selection location associated with the maximum expected gain (MEG), essentially maximizing rewards and minimizing penalties by accounting for variance in end points. People can make rapid estimates of the expected outcomes of their actions in new task configurations (Trommershäuser, Maloney & Landy, 2006). The two sources of information in these decisions are the sensorimotor uncertainty associated with perceiving and acting upon a target as well as the expected gains associated with acting upon the target. Therefore, when making movements under uncertain conditions, people can combine sensory, motor, and reward information to plan adaptive movements.

Such target estimation tasks have captured the mechanisms of single target estimation, but do not describe how such computations would guide decisions between multiple targets. A decision between multiple alternatives might require accumulation and comparison of evidence for multiple targets. A popular model for describing decisions based on time-dependent

accumulation of evidence has been the drift-diffusion model (DDM). DDMs, so-called accumulation-to-bound models, are a useful for describing the psychophysical and neural components of decision making under uncertainty. Modifications of the simple DDM have been used to successfully model neurological and psychological phenomena, such as broad as averaged levels of LIP neuron activity to changes in decisions after a movement is initiated (Gold & Shadlen, 2007; Resulaj et al., 2009). The DDM is an application of signal-detection theory. The process is simple: noisy evidence is accumulated until the evidence exceeds a decision boundary. Then, the corresponding action is initiated. This structure focuses on both the decision made and the reaction time to reach the decision. The process modeled contains parameters such as: drift-rate, which is the overall bias in the evidence accumulated, threshold, which is the height of the decision bounds, and the initial bias, which is how close the starting point is to one choice compared to the other. According to the Believer-Skeptic framework, the competitive dynamics of a neural network model of cortico-basal ganglia pathways should map onto the drift-rate parameter of a DDM, which intuitively expresses the competition and uncertainty associated with decisions. In contrast, shifts in the decision boundary or threshold are more likely to reflect downstream dynamics, such as thalamic or cortical pathways (Pouget et al. 2011).

To investigate how sensory uncertainty and reward properties collectively drive risky decisions, a motor gamble task with perceptually uncertain targets was employed. To explore how these decisions map onto the Believer-Skeptic framework, drift-diffusion models were fit to subject choices and reaction times. If the choice between two perceptually uncertain targets with economic properties can be explained by Believer-Skeptic framework, then increasing the reward magnitude and reward probability of a risky target while increasing the perceptual

uncertainty of a safe alternative should all contribute to an increase in the drift-rate towards the risky target.

METHODS

Participants

All experimental procedures were approved by the institutional review board (IRB) at Carnegie Mellon University (CMU). Supplemental approval for Experiment 1 was obtained from University of Pittsburgh IRB. Twenty participants for Experiment 1 (12 male, 20.9 = mean age) were recruited by flyer and Pittsburgh online subject pool, while participants for Experiment 2 (N = 20, 7 male, 18.8 = mean age) and 3 (N = 20, 11 male, 19.6 = mean age) were recruited by CMU's Psychology Department Student Pool. All subjects gave informed consent prior to participation in the study.

Experimental Design

All experiments were two-alternative forced choice/estimation tasks on a computer and spanned approximately forty-five minutes. The goal of the task was to maximize the amount of points accrue by 1) choosing a target and 2) selecting as close to center of the target using a mouse. In every trial, participants chose between two targets (Figure 2). Targets were colored clouds of dots, whose locations were drawn from a Gaussian surrounding an invisible mean. Targets appeared in a random location on the screen, though the target means were always horizontally adjacent, with the distance between drawn from a uniform distribution of 100-200 pixels. Stimuli were drawn using PsychToolbox3 (MATLAB).

The two targets were distinguishable by their color: the "Risky Target" was consistently indicated by a blue dot cloud, while the "Safe Target" was a yellow dot cloud. The two targets

differed throughout the experiments in their reward properties (magnitude and probability) and their spatial variance, the variance of the 2D Gaussian from which they are drawn. Selecting the safe target guaranteed points, however, in some experiments its variance was varied by trial. The risky target was a gamble with a probability of giving a reward or penalty; the magnitude of rewards and penalties are manipulated in some experiments. The selected target was defined as the target whose mean was closest to the participant's endpoint in a trial, and trial feedback was calculated as the target reward magnitude divided by the distance between the selection and the nearest target mean.

All experiments were split into eight blocks of approximately 100 selection trials. At the beginning of each block, a display provided the following information about the upcoming trials: Risky Target reward probability, maximum Risky Target reward magnitude, and maximum Risky Target penalty magnitude. Since the magnitude of the Safe Target never changed, it was used as a baseline and its value was implicit after the participants were told its value in the introduction to the experiment.

In a single trial of an experiment, the participant went through a standardized procedure (Figure 2). The subject initiated the trial by clicking the mouse. Then, a crosshair at the center of the screen appeared for a jittered period of 250-2500ms. The crosshair disappeared, and the targets flashed simultaneously onto the screen for 300ms. The targets disappeared, and a cross appeared at the center of the screen. Participants moved the cross to their selection location and terminated the trial by releasing the left mouse button. Following the selection, a score indicating performance on the trial was displayed.

All experiments followed this same paradigm, but manipulated different experimental parameters impacting the gambles. Risky Target score magnitudes are represented relative to the

Safe Target magnitude and are presented as ratios of reward-to-penalty (e.g. a 2:-1 reward-to-penalty magnitude indicates that the Risky Target reward was twice as large as the Safe Target, while the penalty was the negative scaling of the Safe Target). Spatial variances of the Safe Target are represented as multiples of the Risky Target spatial variance which was always 33 pixels (e.g. Safe Target spatial variance of 4x indicates 4 times the Risky Target variance, or 132 pixels).

In Experiment 1 (N = 20) we manipulated the Risky Target reward and penalty magnitudes (1:-2, 1:-1, 2:-1, and 4:-1) by block and Safe Target spatial variance (1x, 2x, 4x, and 6x) by trial. In Experiment 2 (N = 20) we manipulated the Risky Target reward and penalty magnitudes (1:-1, 2:-1) and reward probabilities of the Risky Target (0.15, 0.3, 0.75, and 0.85), both by block. In Experiment 3 (N = 20) we manipulated the Risky Target reward probabilities (0.15, 0.3, 0.75, and 0.8) by block and the Safe Target spatial variance (1x, 2x, 4x, and 6x) by trial. Each experiment contained a total of eight blocks with one repeat block for each condition in Experiment 1 and 3. Each block contained 96, 98, and 100 trials for each experiment, respectively.

Modeling

Decisions and reaction times were modeled using the Hierarchical Drift Diffusion Model (HDDM) package, an implementation of Hierarchical Bayesian (parameter) estimation of the Drift-Diffusion Model in Python (Wiecki, Softer, and Frank, 2013). This toolbox allows simultaneous estimation of group and individual subject DDM parameters using Bayesian inference. The parameters are estimated hierarchically by using the group distribution of parameter estimates as a prior to inform estimates of individual subject parameters in creating the model.

For each experiment, a model whereby a unique drift-rate is estimated for each combination of manipulated experimental parameters is fit. Another model where the bias is estimated in the same manner is used as a comparison model. Estimation of the drift-rate and bias parameters was accomplished using five-thousand MCMC samples to construct the Markov chain. For all experiments, positive drift-rates indicate faster evidence accumulation for risky decisions, whereas negative drift-rates indicate biased evidence accumulation for safe decisions. The first two-hundred samples were burned to sample from convergent distributions. Model convergence was assessed by visual inspection of posterior distribution, traces, and autocorrelation. Model convergence is important for interpreting the parameter estimates as stable. Model fit was assessed by the quantitate measure of deviance information criterion (DIC), a hierarchical model's substitute for AIC and BIC (Spiegelhalter, Best, and Carlin, 1998). Lower values of DIC indicate better fits with large, negative values being most desirable. The estimated drift-rates were qualitatively assessed by comparison with subject decisions.

Data Analysis

Subjects' choice behaviors and the model estimate drift-rate parameters were analyzed using two-factor, within-subjects repeated measures, implemented in MATLAB. Custom scripts were developed to process the data for input into the 'rm_anova2' script developed by Aaron Schurger (available at MathWorks). All python and MATLAB scripts and data for this thesis will be made available on the CoAxLab Github (https://github.com/CoAxLab).

RESULTS

In all experiments, a participant's data was only analyzed if less than 5% of their trials had a movement time of 0 seconds. This criterion was chosen to remove subjects who simply

clicked through the experiment without making selections or who had enough error trials (accidental clicks) to warrant exclusion from analysis. From Experiments 1-3, one, four and two subjects, respectively, were excluded from analysis based on this criterion.

Experiment 1

ANOVA of subjects' proportion of risky choices shows the main effects: higher reward-to-penalty magnitude increases risky decision ($F(3,54) = 42.83$, $p < 0.001$), higher safe target spatial variance increases risky decisions ($F(3,54) = 24.94$, $p < 0.001$), and an interaction between magnitude and variance ($F(9,162) = 4.651$, $p < 0.001$) (Table 1; Figure 3).

The drift-rate model determined drift-rates individually based on the Risky Target reward-to-penalty magnitude and Safe Target spatial variance conditions (Figure 3). Threshold, drift-rate, and standard deviation posteriors converged. However, the non-decision time posterior did not converge, even after accounting for inter-trial variability. The model's fit was robust (DIC = - 3,054) and better than that of the bias model (DIC = -1,547) (Table2). Additionally, the drift-rate estimates follow the same trend as subject choices: drift-rate increases with higher Risky Target reward-to-penalty magnitudes and higher Safe Target spatial variance.

ANOVA of subjects' estimated drift-rate parameters shows the main effects: higher reward-to-penalty magnitude increases risky drift-rate ($F(3,54) = 43.71$, $p < 0.001$), higher safe target spatial variance increases risky decisions ($F(3,54) = 49.17$, $p < 0.001$), and an interaction between magnitude and variance ($F(9,162) = 2.738$, $p = 0.0053$) (Table 1; Figure 3).

Experiment 2

ANOVA of subjects' proportion of risky choices shows the main effects: higher reward-to-penalty magnitude increases risky decisions ($F(1,15) = 16.54$, $p < 0.001$), higher reward

probability increases risky decisions ($F(3,45) = 79.43$, $p < 0.001$), and an interaction between magnitude and probability ($F(3,45) = 4.194$, $p < 0.01$) (Table 1; Figure 4).

The model determined drift-rates individually based on the Risky Target reward-to-penalty magnitude and spatial variance conditions (Figure 4). Threshold, drift-rate, standard deviation, and non-decision time posteriors converged successfully. The model's fit was robust (DIC = -25,471) and better than that of the bias model (DIC = -18,796) (Table2). Additionally, the drift-rate estimates follow the same trend as subject choices: drift-rate increases with higher Risky Target reward-to-penalty magnitudes and higher Risky Target reward probability.

ANOVA of subjects' estimated drift-rate parameters shows the main effects: higher reward-to-penalty magnitude increases risky drift-rate ($F(1,15) = 11.84$, $p < 0.005$), higher reward probability increases risky drift-rate ($F(3,45) = 71.19$, $p < 0.001$). However, an interaction between magnitude and probability could not be detected ($F(3,45) = 2.530$, $p = 0.069$) (Table 1; Figure 4).

Experiment 3

ANOVA of subjects' proportion of risky choices shows the main effects: higher reward probability increases risky decisions ($F(3,51) = 33.41$, $p < 0.001$), higher safe target spatial variance increases risky decisions ($F(3,51) = 146.70$, $p < 0.001$), and an interaction between probability and variance ($F(9,153) = 4.601$, $p < 0.001$) (Table 1; Figure 5).

The model determined drift-rates individually based on the Risky Target reward-to-penalty magnitude and spatial variance conditions (Figure 5). Threshold, drift-rate, standard deviation, and non-decision time posteriors converged successfully. The model's fit was robust (DIC = -23,947) and better than that of the bias model (DIC = -18,692) (Table2). Additionally,

the drift-rate estimates follow the same trend as subject choices: drift-rate increases with higher Risky Target reward-to-penalty magnitudes and higher Risky Target reward probability.

ANOVA of subjects' estimated drift-rate parameters shows the main effects: higher reward probability increases risky drift-rate ($F(3,54) = 90.92$, $p < 0.001$), higher safe target spatial variance increases risky drift-rate ($F(3,54) = 41.66$, $p < 0.001$), and an interaction between probability and variance ($F(9,162) = 2.825$, $p < 0.005$) (Table 1; Figure 5).

DISCUSSION

In each of these experiments, perceptual uncertainty, reward magnitudes, and reward probabilities impacted decision preferences. It is shown here that perceptual uncertainty does influence value-based decisions; when the perceptual uncertainty of a risky choice is low relative to a safe alternative, people make more risky choices regardless of the Risky Target's value. In accordance with the Believer-Skeptic framework, this effect is best modeled as an increase in the drift-rate towards the Risky Target. The reward magnitude and reward probability similarly best mapped onto the drift-rate towards the Risky Target.

In response to manipulations of reward-to-penalty magnitudes of the Risky Target, participants universally selected the Risky Target more often when it's scaling favored larger rewards relative to penalties. It should be noted that rewards and penalties are not weighted equally. The zero-risk bias is the phenomenon where a target with no risk is preferred to an equivalently valued target with risk. This effect is observed in Experiment 1 by observing the preference for the Risky Target when its reward-to-penalty magnitude is 2: -1 and the Safe Target spatial variance is equivalent to the Risky Target's variance (Figure 3). An unbiased gambler would choose each target with equal probability; we observe a substantial bias for the

Safe Target, which was selected 75% of the time for these trials. This behavior was one of the first hints that humans are not strictly rational decision-makers in the classical economic utility sense.

Additionally, participants adapted decision-making to the probability of receiving rewards from the Risky Target. Looking to Experiment 2, we see that participants exhibit non-linear weighting of reward probability. If the probability was weighted linearly, we would have seen a consistent change in choice for each percent change in reward probability. In cumulative prospect theory, this is referred to as persons' weighted probability functions (Tversky & Kahneman, 1992). Generally, these curves are better fit by logistic rather than linear regressions. Such comparisons between economic and motor gambles have been conducted and investigated these curves; it was shown that people tended to be more risk seeking in their motor gambles compared to equivalent classic economic gambles (Wu, Delgado, and Maloney, 2009). In motor gambles, there is a tendency to underweight low probabilities and overweight high probabilities. In Experiment 2, we show that this relationship is modulated by the reward and penalty magnitudes associated with the gamble since the change in choice preference is not the same between probabilities for the two reward values used. This experiment is done with moderate probabilities, however, and we did not explore how people weight the extremes (less than 0.1 and greater than 0.9). If we had, we would expect to see the same tendency to underweight low probabilities and overweight the high probabilities.

Finally, spatial variance of the Safe Target universally increased the rate at which participants chose the Risky Target. It is known that increasing the spatial variance increases the variance of selections (Faisal, Selena, & Wolpert 2008; Jarbo, Flemming, and Verstynen, in review). In this manner, the spatial variance directly impacts the reward feedback from the Safe

Target. In experiments 1 and 3, we see that participants choose the Risky Target more frequently when the spatial variance of the Safe Target is high. One additional way in which spatial variance may decrease the value of the Safe Target is by increasing the challenge of estimating the center of the target. This effect would be dissociable from the effect on reward feedback. However, in this experiment, the reward errors and sensory errors are not easily dissociated. So, determining the degree to which spatial variance influences decision-making because of payoff versus computational challenge would require further investigation.

The Believer-Skeptic framework allows us to interpret the biological mechanisms behind these contextual influences on risky decision-making. First, the two-alternative forced choice drift-rates estimated here reflect the difference in evidence accumulation between the two targets. The evidence accumulated for a decision is the algorithmic description of the dynamic competition between the direct and indirect pathways in the basal ganglia (Dunovan & Verstynen, 2016; Figure 1). These two pathways are described as mutually inhibiting, recurrently exciting neural networks. The direct pathway population for the Risky Target receives input about the value of the Risky Target, while the indirect population acting on the Risky Target receives input from the value of alternative choices (i.e. the Safe Target). The believer (direct) population competes to convince the skeptic (indirect) population of the value of the Risky Target, while the same process occurs among another population for the Safe Target. This process naturally captures uncertainty associated with the two targets. Importantly, this framework predicts that the information is encoded in the drift-rate of a drift-diffusion model. The results of this study directly agree with this prediction. Therefore, the actions made by our participants likely reflects the population activity of direct vs. indirect pathways involved in their decisions.

In this study, two models that map task parameters (reward-to-penalty magnitude, reward probability, and spatial variance) to DDM (drift-rate and bias) parameters were tested. We observed that a model in which drift-rate is determined by the combination of task parameters fit better than a model estimating the bias parameter in the same manner. However, we did not test other parameter mappings, or combinations of mappings. For example, drift-rate determined by reward-to-penalty magnitudes but bias determined by variance. These alternatives were tested in the proposition of the Believer-Skeptic framework (Dunovan & Verstynen, 2016). If we were to compare the drift-rate model with a more extensive list of possible alternatives and it still turned out the best fit, we could be more confident in drawing conclusions based on the Believer-Skeptic framework.

Upon debriefing, many participants did indicate some explicit strategies for approaching these tasks. For example, some participants stated choosing targets based on the targets' distances from the center of the screen, supposedly in optimization of Fitts's Law (Fitts, 1992). Another common statement by participants was that they had decided to select only one target given a certain condition (e.g. high reward probability, or high spatial uncertainty). Others indicated explicit strategies similar to a win-stay, lose-switch heuristic. It would be interesting to investigate whether it would be possible to capture the use of such strategies in different portions of the tasks. Since we might expect these cognitive processes to be captured by the DDM, it would be interesting to split the task into chunks based on where subjects might have implemented strategies, and to determine whether each section of the task is still modeled best by the drift-rate model. A more thorough approach would be to design an experiment that requires either the accumulation of evidence or the implementation of a heuristic to make a decision. We would expect that the parameter mappings would be different given the two scenarios. Despite

the fact that our only comparison was the bias model, the drift-rate model implemented in this study is robust. Therefore, a mapping whereby value and spatial uncertainty of a gamble determines the rate of evidence accumulation is an effective cognitive description of perceptual motor gambles. Since this observation agrees with the Believer-Skeptic framework, we can postulate that economic and sensory information in value-based sensorimotor decisions are combined through the competitive dynamics of the direct and indirect basal-ganglia pathways associated with each action.

References

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience,13*(01), 87-100. doi:10.1017/s095252380000715x

Cognitive Axon Laboratory. (n.d.). Retrieved from https://github.com/CoAxLab

Dunovan, K. (2017). A biologically motivated synthesis of accumulator and reinforcement-learning models for describing adaptive decision-making (Doctoral dissertation, University of Pittsburgh)

Dunovan, K., & Verstynen, T. (2016). Believer-Skeptic Meets Actor-Critic: Rethinking the Role of Basal Ganglia Pathways during Decision-Making and Reinforcement Learning. *Frontiers in Neuroscience,10*. doi:10.3389/fnins.2016.00106

Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience,9(4), 292-303*. doi:10.1038/nrn2258

Fitts, P. M. (1992). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology: General,121*(3), 262-269. doi:10.1037//0096-3445.121.3.262

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience,30*(1), 535-574. doi:10.1146/annurev.neuro.29.051605.113038

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature,394*(6695), 780-784. doi:10.1038/29528

Jarbo, K., Flemming, R., & Verstynen, T. D. (2017, January 17). Sensory uncertainty impacts avoidance during spatial decisions. Retrieved from osf.io/preprints/psyarxiv/xgpfk

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature,427*(6971), 244-247. doi:10.1038/nature02169

Mink, J. W. (1996). The Basal Ganglia: Focused Selection And Inhibition Of Competing Motor Programs. *Progress in Neurobiology,50*(4), 381-425. doi:10.1016/s0301-0082(96)00042-1Pouget, P., Logan, G. D., Palmeri, T. J., Boucher, L., Pare, M., & Schall, J. D. (2011). Neural Basis of Adaptive Response Time Adjustment during Saccade Countermanding. *Journal of Neuroscience,31*(35), 12604-12612. doi:10.1523/jneurosci.1868-11.2011

Spiegelhalter, D. J., Best, N. G., & Carlin, B. P. (1998, March 30). Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models. Retrieved from http://s3.amazonaws.com/academia.edu.documents/40156925/Bayesian_Deviance_the_Effective_Number_o20151118-18894-1vfafh9.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1490550578&Signature=X8HGaMEsACRY9d7Mwt9Yctj0Cv8%3D&response-content-disposition=inline%3B%20filename%3DBayesian_deviance_the_effective_number_o.pdf

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A,20*(7), 1419. doi:10.1364/josaa.20.001419

Trommershäuser, J., Landy, M. S., & Maloney, L. T. (2006). Humans Rapidly Estimate Expected Gain in Movement Planning. *Psychological Science,17*(11), 981-988. doi:10.1111/j.1467-9280.2006.01816.x

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty,5*(4), 297-323. doi:10.1007/bf00122574

Wiecki TV, Sofer I and Frank MJ (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. Front. Neuroinform. 7:14. doi: 10.3389/fninf.2013.00014

Wu, S., Delgado, M. R., & Maloney, L. T. (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences,106*(15), 6088-6093. doi:10.1073/pnas.0900102106

TABLES

Table 1

Two-Factor, Within-Subjects Repeated Measures ANOVA Results for Behavioral Choices (Beh) and Drift-Rate Parameters (V) in each Experiment

Experiment 1	Beh	V	*Experiment 3*	Beh	V
Reward: Penalty Magnitudes	$F(3,54) = 42.83$ $p < 0.001$	$F(3,54) = 43.71$ $p < 0.001$	Reward Probability	$F(3,51) = 33.41$ $p < 0.001$	$F(3,54) = 90.92$ $p < 0.001$
Stable Target Variance	$F(3,54) = 24.94$ $p < 0.001$	$F(3,54) = 49.17$ $p < 0.001$	Stable Target Variance	$F(3,51) = 146.70$ $p < 0.001$	$F(3,54) = 41.66$ $p < 0.001$
Magnitude x Variance	$F(9,162) = 4.651$ $p < 0.001$	$F(9,162) = 2.738$ $p = 0.005$	Proability x Variance	$F(9,153) = 4.601$ $p < 0.001$	$F(9,162) = 2.825$ $p < 0.005$
Experiment 2					
Reward: Penalty Magnitudes	$F(1,15) = 16.54$ $p < 0.001$	$F(1,15) = 11.84$ $p < 0.005$			
Reward Probability	$F(3,45) = 79.43$ $p < 0.001$	$F(3,45) = 71.19$ $p < 0.001$			
Magnitude x Probability	$F(3,45) = 4.194$ $p < 0.01$	$F(3,45) = 2.530$ $p = 0.069$			

Table 2

Deviance Information Criterion (DIC) of Drift-Rate and Bias Models for Each Experiment

Model	Experiment 1 DIC	Experiment 2 DIC	Experiment 3 DIC
Drift-rate	- 3054	-25,471	-23,947
Bias	-1,547	-18,796	-18,692

FIGURES

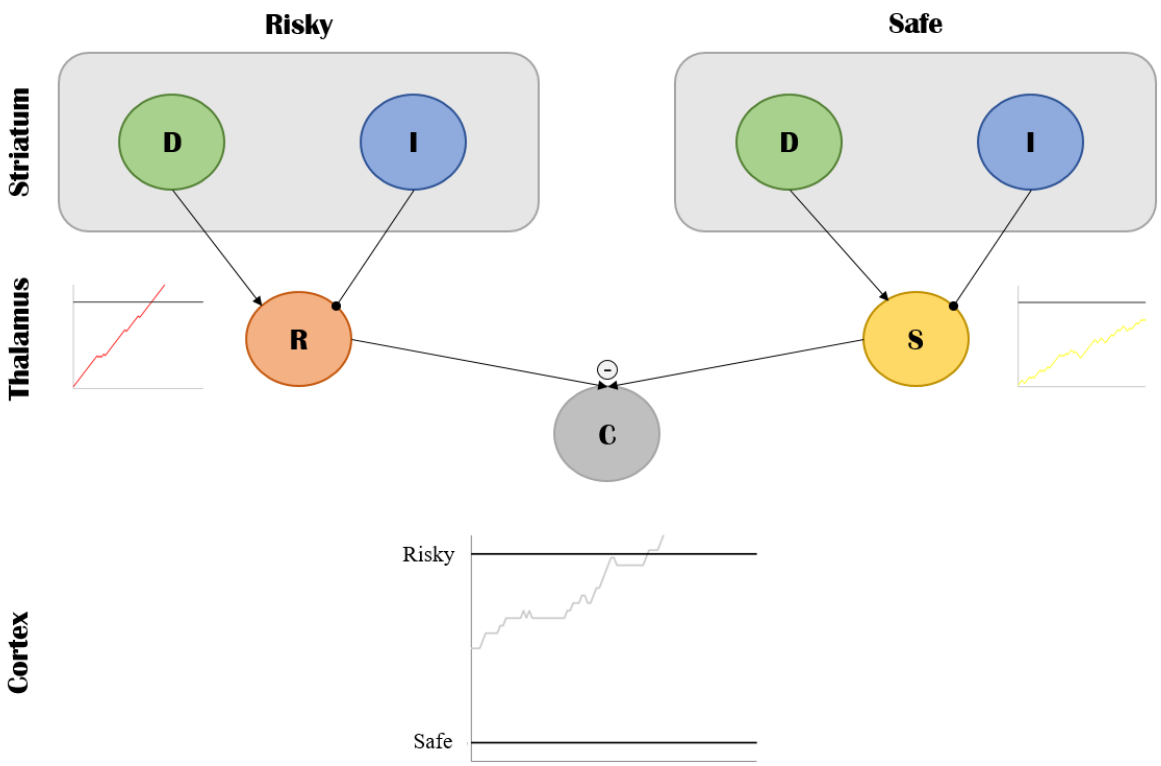

Figure 1. A schematic of the Believer-Skeptic Framework, applied to the experimental design. The direct (D) and indirect (I) pathways compete to activate populations responsible for the risky (R) and the safe (S) decision. The competition between the populations' activities in the cortex (C) can be algorithmically summarized as a DDM with evidence accumulating towards either a risky or safe decision boundary. Arrows indicate excitatory inputs whereas knobs indicate inhibitory inputs.

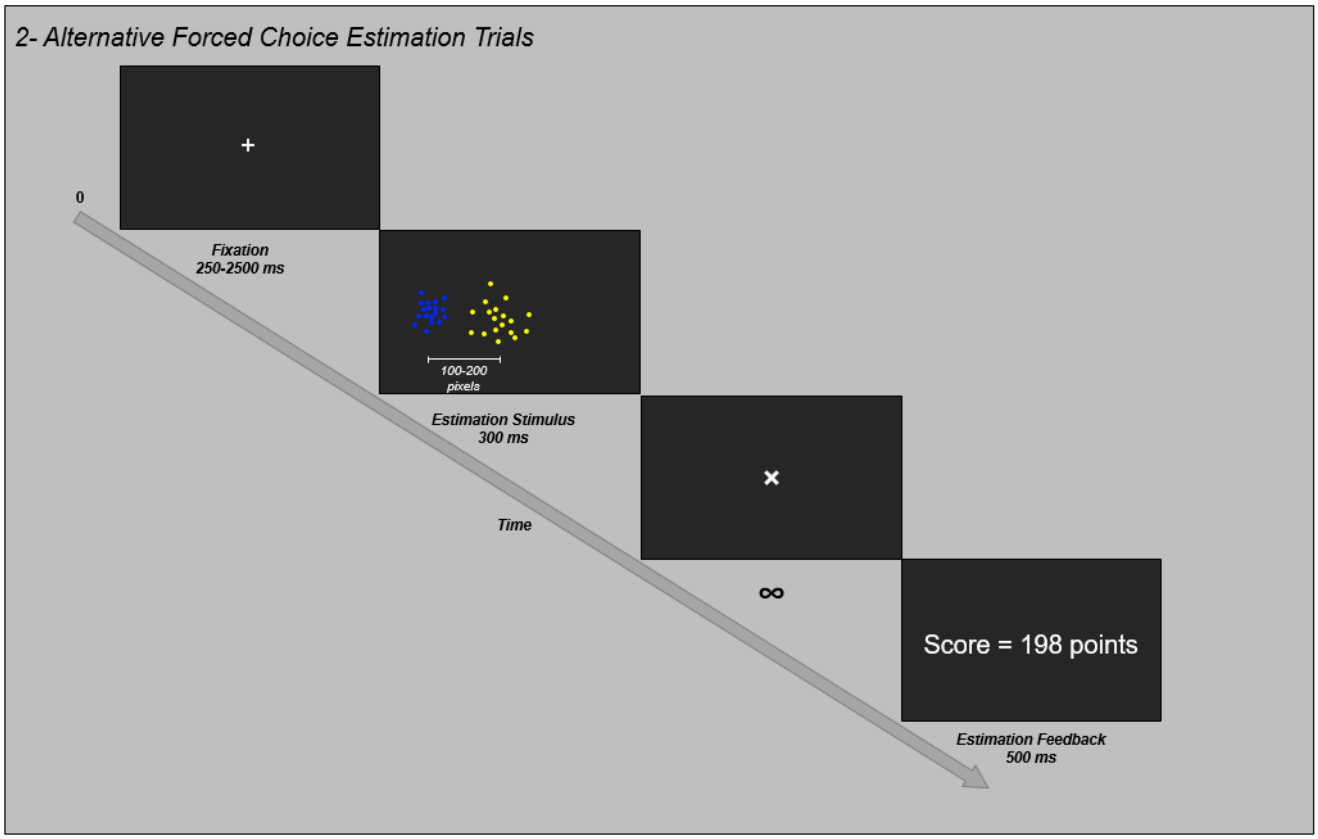

Figure 2. Time-course of a single trial. All Trials are initiated when the subject clicks the mouse. The subject leaves the mouse held until phase 3, where they move the 'x' to their selection location. Following the estimation feedback, the screen goes blank, and the process repeats.

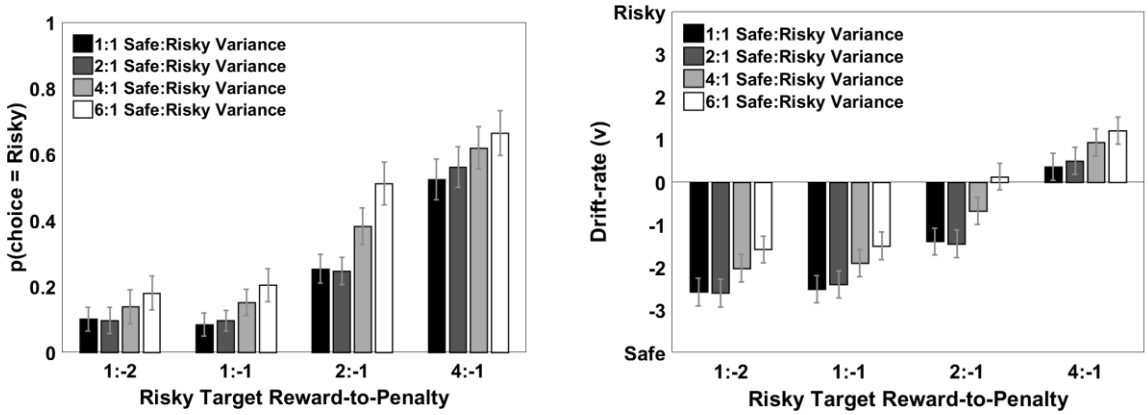

Figure 3. The effects of reward-to-penalty magnitude (groups) and spatial variance (color coded) on decision preferences is reflected in the drift-rate of the DDM fitted to Experiment 1. Increased Risky Target reward and increased Safe Target variance increase risky decisions and the model drift-rate. Spatial variance is coded as the ratio of Safe Target variance to Risky Target variance.

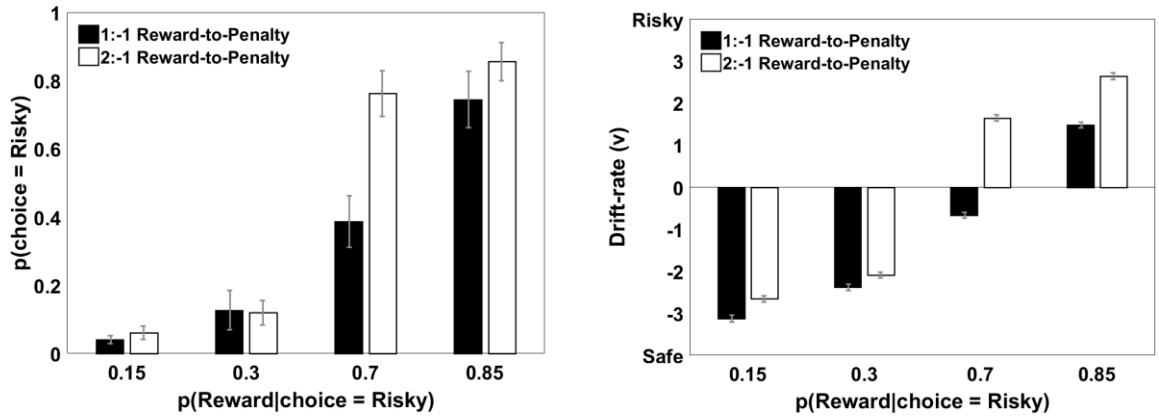

Figure 4. The effects of Risky Target reward probability (groups) and reward-to-penalty magnitudes (color coded) are reflected in both subject group decisions and drift-rates of the DDM. Increased probability of reward form the Risky Target and increased reward-to-penalty magnitudes increase risky decisions and drift-rate.

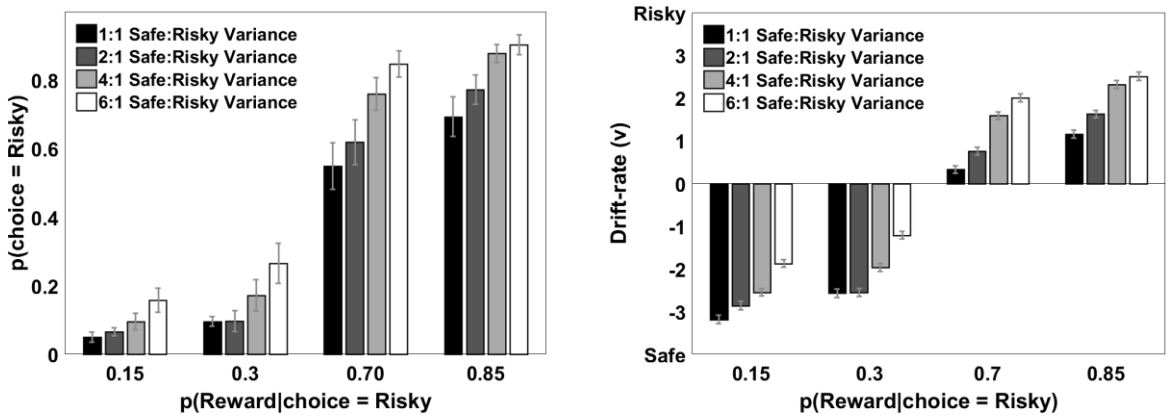

Figure 5. The effects of Risky Target reward probability (groups) and Safe Target spatial variance (color coded) are observed in both subject group decisions and drift-rates of the DDM. Increased Risky Target reward probability and Safe Target spatial variance increase risky decisions and drift-rate.