

# Risk assessment in man and mouse

Fuat Balci<sup>1</sup>, David Freestone<sup>2</sup>, and Charles R. Gallistel<sup>3</sup>

Department of Psychology and Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854-8020

Contributed by Charles R. Gallistel, December 16, 2008 (sent for review October 16, 2008)

**Human and mouse subjects tried to anticipate at which of 2 locations a reward would appear. On a randomly scheduled fraction of the trials, it appeared with a short latency at one location; on the complementary fraction, it appeared after a longer latency at the other location. Subjects of both species accurately assessed the exogenous uncertainty (the probability of a short versus a long trial) and the endogenous uncertainty (from the scalar variability in their estimates of an elapsed duration) to compute the optimal target latency for a switch from the short- to the long-latency location. The optimal latency was arrived at so rapidly that there was no reliably discernible improvement over trials. Under these nonverbal conditions, humans and mice accurately assess risks and behave nearly optimally. That this capacity is well-developed in the mouse opens up the possibility of a genetic approach to the neurobiological mechanisms underlying risk assessment.**

genetics | human | interval timing | optimality | statistical decision theory

Humans and other animals make decisions in the face of uncertainties arising from both exogenous and endogenous stochastic processes. The uncertainty about the outcome of a coin toss is exogenous, whereas the uncertainty in your estimate of the time elapsed since you started reading this paragraph is endogenous. Historically, most of the research conducted in the area of decision-making in the face of risk has focused on the nonnormative phenomena repeatedly demonstrated in decision making that relies on verbally or numerically specified exogenous probabilities (1, 2).

A recent, very different approach to this problem tests subjects' ability to choose the optimal aiming point in a rapid target-touching task in which endogenous motor variability creates risks of either missing the positive pay-off region or straying into a negative pay-off region (3). From the perspective of normative statistical decision theory, the choice of an optimal aiming point is computationally as complex or more complex than the identification of the optimal choice in the verbal tasks. It requires weighing both pay-off magnitudes and probabilities, including very small probabilities at the tails of the distribution of landing points around a target point. Surprisingly, under most conditions, subjects rapidly identify a nearly optimal target point (4, 5). They also optimize the timing of their motor end-points (6, 7).

Inspired by this work, we adapted a simple timing task that requires subjects to judge when it is time to switch from betting on an option with a short-latency payoff to betting on an option with a longer-latency payoff. The subjects were adult humans and a common strain of laboratory mouse (C57/B6). The task was essentially the same for both; they had to anticipate the appearance of a reward at 1 of 2 locations, with often repeated trials. On some fraction of the trials, the reward appeared in one location after a fixed short latency; on the complementary fraction, it appeared at the other location after a fixed longer latency. The trials were randomly intermixed, so subjects did not know at trial outset whether it was a short or long latency. They began each trial by waiting at the short-latency location. If and when they judged that the short latency had passed without a payoff, they switched to the long-latency location. If they switched too soon on a short trial or too late on a long trial, they

lost the reward (or, under some conditions with human subjects, they were assessed penalty points).

The probability of a given outcome depended jointly on the probability of a short versus a long trial and the probability of the subject's switching too soon or too late. Our data are the switch latencies on the long trials, the trials when switching paid-off. Subjects' switch latencies varied from trial-to-trial because of the well-established scalar variability (Weber's law characteristic) in estimates of elapsed duration, which is seen in both human and nonhuman subjects (8–11); repeated estimates of the same duration are accurate only to within roughly  $\pm 15\%$  (12). The fraction of short trials, which varied between blocks of many hundred trials each, was the parameter of the exogenous stochastic process. The variability in a subject's estimate of the time elapsed in a trial and the subject's choice of a target latency for switching were the parameters of the Gaussian endogenous process. From our knowledge of the objective probability of a short trial and an estimate of the standard deviation in a subject's estimate of elapsed duration, it is possible to compute the optimal target time for a switch, that is, the optimal value of the mean switch latency. In our data analysis, we asked how closely our estimates of the optimal switch time,  $\hat{T}_o$ , corresponded to our estimate of the mean of the distribution of a subject's switch latencies,  $\hat{T}$ .

## Results

Fig. 1 shows plot  $\hat{T}_o$ , which is our estimate of the optimal mean switch latency, versus  $\hat{T}$ , which is our estimate of the mean of the distribution from which the sample of observed switch latencies was drawn, for the human (A) and mouse (B) subjects, respectively. We computed the mutual regression between the 2 estimates across different conditions for individual subjects. For human subjects, the average slope was 0.47, which was significantly  $>0$  [ $t(6) = 3.12, P < 0.05$ ] and  $<1$  [ $t(6) = 3.53, P < 0.05$ ]. For mouse subjects, the average slope was 1.31, which was significantly  $>0$  ( $t(11) = 6.15, P < 0.0001$ ) and not significantly different from 1 ( $P = 0.18$ ). Fig. 1C plots group averages of  $\hat{T}_o$  and  $\hat{T}$  as functions of the probability of a short trial in the mouse data.

Fig. 1 A–C show that the observed mean switch latencies closely tracked our estimate of the optimal target latencies. The average absolute temporal distance between the two, ( $\hat{T} - \hat{T}_o$ ), was  $172 \pm \text{SE } 33$  ms in the human subjects, which was 5% of the 3-s range of possible values that  $\hat{T}$  could take. In the mice subjects, the average absolute temporal distance was  $436 \pm \text{SE}$

Author contributions: F.B. designed research; F.B. and D.F. performed research; F.B., D.F., and C.R.G. analyzed data; and F.B. and C.R.G. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

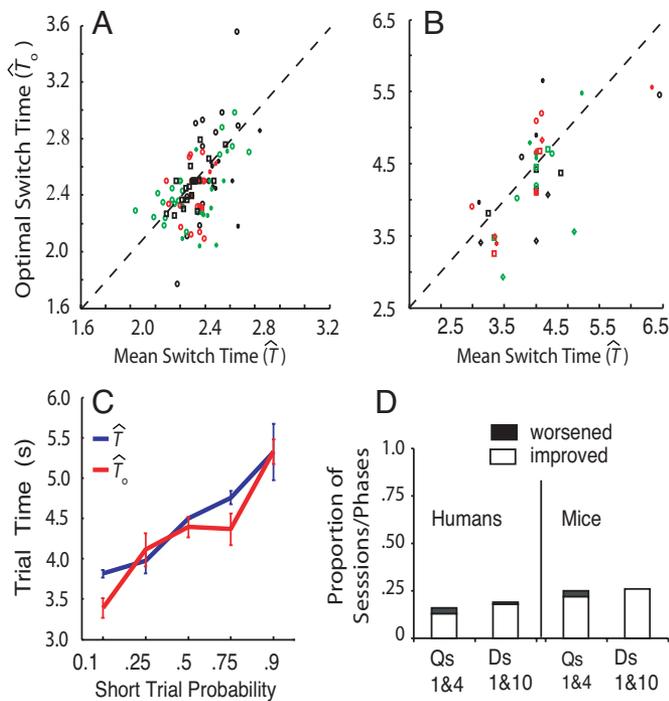
<sup>1</sup>To whom correspondence may be sent at the present address: Princeton Neuroscience Institute, Princeton University, Green Hall, 3-N-12, Princeton, NJ 08544. E-mail: fbaldi@princeton.edu

<sup>2</sup>Present address: Department of Psychology, Brown University, 89 Waterman Street, Box 1853, Providence, RI 02912.

<sup>3</sup>To whom correspondence may be addressed. E-mail: galliste@rucss.rutgers.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0812709106/DCSupplemental](http://www.pnas.org/cgi/content/full/0812709106/DCSupplemental).

© 2009 by The National Academy of Sciences of the USA



**Fig. 1.** Performance of human and mouse subjects. (A and B)  $T_o$  versus  $\hat{T}$  for human and mouse subjects, respectively, with different symbols for different subjects. (C)  $\hat{T}$  and  $\hat{T}_o$  as functions of the probability of a short trial in mouse subjects (means  $\pm 1$  SE). (D) The proportion of experimental conditions (sessions in humans or experimental phases in mice) in which performance improved significantly (white bar) or worsened significantly (black bar) between the first and last quartiles or first and last deciles in the sequence of trials within a given condition.

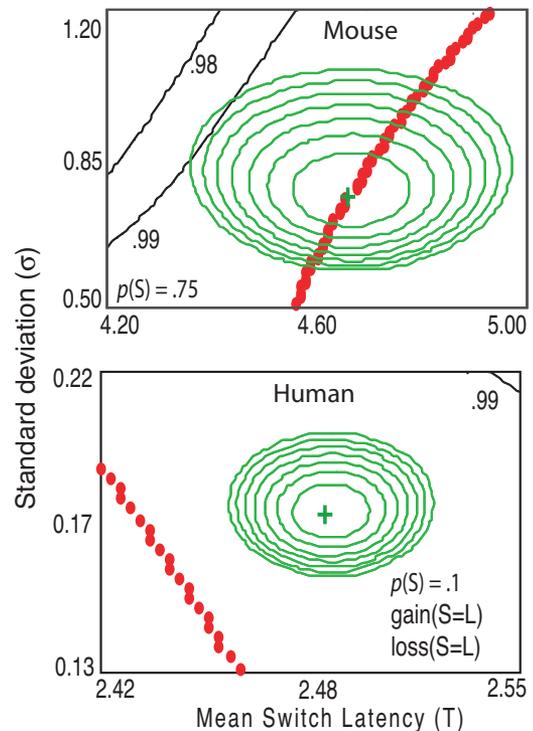
60 ms, which was 7% of the 6-s range of possible values that  $\hat{T}$  could take.

These absolute temporal differences are not meaningful until we take into account the statistical uncertainties surrounding the estimates from which they derive.  $\hat{T}$ , the mean of a sample of switch latencies, is only an estimate of the mean of the population from which the sample was drawn. Similarly, our values for  $\hat{T}_o$  depend on the within-sample variability, from which an estimate of the population SD derives. To take into account the statistical uncertainty about the true values of the population parameters,  $T$  and  $\sigma$ , we plotted for each subject in each condition the likelihood function and the relative expected-gain function on the plane defined by the possible values of  $T$  and  $\sigma$  (Fig. 2).

Assuming an uninformative prior, the likelihood function tells us how likely it is that any given parameter pair,  $(T_i, \sigma_i)$ , are the true values for the mean and variability of the subjects' underlying switch-latency distribution, given the data in our sample from this distribution. The source distributions were assumed normal, because most of the datasets were better fit by normal distributions than by log normal or Weibull distributions.

The expected gain at a point in the parameter plane is the expected gain for a decision maker with known variability  $\sigma_i$  in his/her representation of elapsed time, by using decision criterion  $T_i$ . The "relative" expected gain is this expected gain normalized by the expected gain for a decision maker with the same temporal uncertainty ( $\sigma_i$ ), by using the "optimal" decision criterion (target switch time).

With both the relative likelihood function for  $T$  and  $\sigma$  and the relative expected-gain function mapped onto the common parameter plane, the subject's optimality may be measured either by the relative likelihood of the closest point on the ridge of the



**Fig. 2.** Representative plots of level curves of the likelihood function (green ovals), the optimal mean switch latency (red dots), and relative gain contours (black curves) under varying conditions:  $p(S)$  = probability of a short trial; gain and loss, humans only, is the relation between points won (gain) or lost given the 4 possible outcomes on a trial.

gain function (the subject's statistical distance from the optimal point) or by relative expected gain at  $(\hat{T}, \hat{\sigma})$ , which is the mode of the likelihood function, the point marked by a "+" in the innermost ovals in Fig. 2. This mode is the best estimate of the population parameters.

A second measure is the estimate of the reduction in expected gain that the subject incurs by using a nonoptimal target latency. From the subject's perspective, it is the more appropriate measure. A subject whose expected gain is 99% of the gain to be expected from an optimal choice of target latency would rationally be indifferent to the possibility that their target latency was "very significantly" different from the optimal target latency.

Fig. 2 shows level curves (equal-likelihood contours) of the likelihood function, together with contours of the relative expected gain function, for 2 different conditions (one with a human subject and one with a mouse subject). Fig. 2 also shows the locus of the optimal decision criterion (mean switch latency) as a function of the variability in the switch latencies. The relative-likelihood contours are at natural log unit levels ( $e^{-1}$ ,  $e^{-2}$ , etc). If the innermost relative likelihood contour encompasses a red dot (a point on the ridge of the expected gain function), then the statistical distance from optimality is negligible. Conversely, if the outermost relative likelihood contour does not encompass a red dot, then the difference between the subject's estimated decision criterion and the optimal criterion is, from a purely statistical perspective, highly significant. Whether this statistical difference is of any practical consequence depends on the relative expected gain at the "+" in the center of the innermost relative likelihood contour. When the "+" is inside the 0.99 relative gain contour, the best estimate of the subject's expected gain is >99% of the maximum possible expected gain for a subject with that degree of variability in their switch latencies (hereafter referred to as relative score). In such cases,

the statistical distance from the optimal, whether large or small, is of no consequence.

Human subjects' median relative score was 0.98 (interquartile interval: 0.02), and mice subjects' relative score was 0.99 [interquartile interval: 0.01 (excluding from the human data conditions where the maximum possible expected gain was  $\leq 0$ )]. Thus, we conclude that both humans and mice can compute a temporal decision criterion—a target switch time—that is sufficiently close to the optimal so as to very nearly maximize their expected gain.

Subjects did not appear to converge on a nearly optimal decision criterion by trial-and-error hill-climbing processes like those proposed by operant psychologists (13, 14) and advocated within the reinforcement-learning tradition in computer science (15). If they did, we should consistently observe that  $|\hat{T}_o - \hat{T}|$ , the difference between our estimate of a subject's decision criterion and our estimate of the optimal criterion, grew smaller as the number of trials under a given condition increased. However, in the great majority of cases, there was no discernible improvement in the location of a subject's target switch latency over trials (Fig. 1D).

## Discussion

In this simple timing task, which we believe captures the essence of temporal decision making that confronts human and nonhuman animal subjects in everyday life, both humans and mice are nearly optimal decision makers in the face of uncertainties of both external and internal origin. This nearly optimal performance would seem to imply that they rapidly form an accurate representation of both their endogenous uncertainty (how good they are) (cf. 16) and the exogenous uncertainty (the objective risk) and that they combine these uncertainty estimates in a normative way. The nearly optimal performance of humans in our task contrasts with the traditional view that humans are nonnormative decision-makers under probabilistic conditions.

Given our results it can be argued that nonrational decisions made under probabilistic conditions in traditional decision-making tasks derive in substantial measure from the form in which the probabilistic information is presented to the subjects (see also refs. 17–19). Specifically, in traditional paper–pencil decision-making tasks, subjects are presented with the exogenous probabilistic information via one-shot verbal descriptions and asked to combine this information with other probabilistic information and/or magnitudes. By contrast, in our task where subjects performed near optimally, they experience both the exogenous and endogenous uncertainty, and they must combine these representations of uncertainty and magnitudes to arrive at an estimate of an appropriate duration target.

We conclude that the decision-making processes might take normative account of probabilistic input only when the probabilities are derived from direct experience of the generative stochastic product and that language-derived representations of uncertainty may not be able to access these processes. Alternatively, the linguistic interface may lead to systematic distortions of the specified uncertainties (e.g., overestimation of the low probabilities).

In our experiments, the output of the decision-making process in both humans and mice was instantiated by temporally controlled responses. Given the ubiquity of timing as a cognitive domain (ranging from fish to humans), the use of a temporal discrimination paradigm (thus temporal information processing) as the context for decision-making under uncertainty adds a translational character to our decision-making task. The translational nature of this cognitive domain has started to receive increasing appreciation in preclinical research (20–21).

The similarity of the results from human and mice subjects in this complex temporal decision-making task further supplies the behaviorally necessary tools to study the genetic and neurobio-

logical underpinnings of risk assessment in animal models. Risk is not an exotic aspect of a uniquely human life; risk is inherent in life at every level of complexity. Thus, it is not far-fetched to suppose that mechanisms for near-optimal risk assessment in many everyday contexts evolved long ago (cf. 19). An obvious question is whether there may exist ways of tapping into those mechanisms to mediate the many human assessments of risk that depend on verbal or symbolically transmitted information about probabilities.

## Methods

**Subjects. Human.** The subjects were undergraduate students, graduate students, and postdoctoral volunteers at Rutgers University. There were 7 subjects, 4 males and 3 females. The age of subjects ranged between 21–34. They gave informed consent and were paid for their participation.

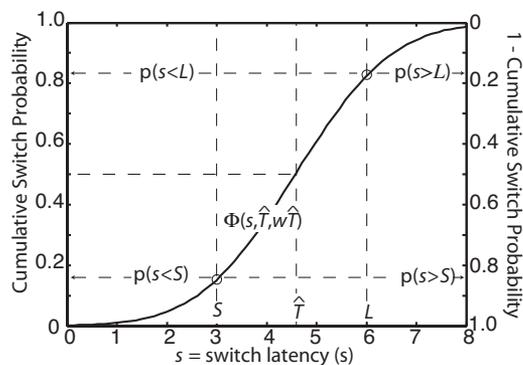
**Mice.** Twelve naive C57BL/6N female mice (Harlan, Indianapolis, IN) were used in this experiment. They were 8 weeks old on arrival. The mice were housed individually in polypropylene cages. The cabinets that held the cages were lit on a 12:12 light/dark cycle (lights on at 8:00 PM). The experiments were run during the dark cycle. The mice were maintained at 85% of their free-feeding weight by being fed lab chow after each session. During the sessions, they were fed 20-mg Noyes food pellets (PJA1-0020) as a reinforcer. Water was available ad libitum in the home cages and experimental chambers. They were treated in accordance with the *Guide for the Animal Care and Use of Laboratory Animals* (22).

**Apparatus. Human.** The temporal parameters, presentation of the stimulus, and recording of the responses were controlled with a Macintosh running OS X 10.3.9. The controlling program was written in MATLAB 7.04 by using the Psychophysics Toolbox extensions (23).

**Mice.** Six operant chambers (Med Associates, ENV 307-W: 21.6 cm  $\times$  17.8 cm  $\times$  12.7 cm) inside ventilated, sound-attenuated boxes (Med Associates, ENV-018-M: 55.9 cm  $\times$  55.9 cm  $\times$  35.6 cm) were used in this experiment. Two opposing sidewalls were made of metal, the other 2 walls and the ceiling of clear Plexiglas. Each chamber had a grid floor with 25 evenly spaced metal rungs. Two pellet dispensers (ENV-203–20) were able to deliver food to 2 cubic hoppers (Med Associates ENV-203-20, 24 mm on each side) spaced 5 inches apart on one wall. On the opposite wall, another cubic hopper was used to initiate trials. Along that same wall, a water bottle protruded into the chamber. The hoppers were illuminable and equipped with an infrared beam that detected nose pokes. A white-noise generator delivered an 80-dB, flat 10–25,000-Hz white noise for programmable durations. The experiment was controlled by software (Med-PC IV, Med Associates) that logged and time-stamped the events—the onsets and offsets of interruptions of the IR beams in the station, the onsets and offsets of white noise, and the delivery of food pellets. Event times were recorded with a resolution of 20 ms.

**Procedure General Procedure.** There were 2 types of trials. On short trials, a signal (a visual stimulus for humans and visual/auditory stimuli for mice) appeared and stayed on for a short latency. The subjects were rewarded if they made the response associated with the short latency at (or after, for mice) the termination of the signal. On long trials, the same signal appeared, but stayed on for a long latency, and the subjects were rewarded if they made the response associated with the long latency at (or after, for mice) the termination of the signal. The response for the long latency was in a different location than the response for the short latency. All subjects started their responding in the short latency location and switched to the long latency location if and when they judged that the trial was a long trial. If the subject emitted the long latency response at (or after, for mice) the termination of the signal in a short trial (e.g., early switch) or the short response at (or after, for mice) the termination of the signal in a long trial (e.g., late switch), it was a missed trial. There were 2 parameters which were varied, in the expectation that they would have an effect on the mean switch latency: (i) The probability of a short trial and (ii) the Payoff Matrix; the rewards for catching short or long latency targets and penalties for switching early in short-latency trials and late in long-latency trials. The latter factor was varied only with human subjects.

**Humans.** The short latency was 2 seconds, and the long latency was 3 seconds. A cross (+) appeared in the middle of the computer screen to start a trial. A press on the "V" key caused an open red square to appear to the left of the +. The "B" key caused it to appear to the right. The open square did not stay (was not effective in catching the target) after the key was released. Thus, the subject was required to keep the key pressed at all times during the trial.



**Fig. 3.** The cumulative distribution of switch times,  $\Phi(s, \hat{T}, w\hat{T})$ , determines the probabilities of premature departure,  $p(s < S)$ ; on-time departure,  $p(s > S)$ ; in-time departure,  $p(s < L)$ ; and too-late departure,  $p(s > L)$ . In-time departures are rewarded; whether a premature or too-late departure is rewarded depends on the complementary probabilities of the trial being short or long.

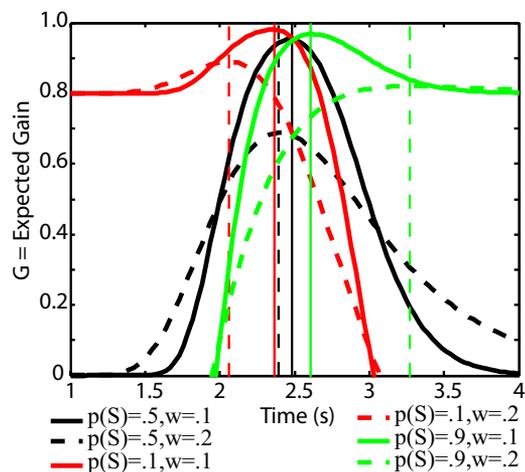
Whether the trial was short or long, if the open square was in the correct location at the time of signal termination, points were added to their score. There were 2 types of errors: A premature switch (before the short trial interval had elapsed) and a late switch (after the long trial interval had elapsed). Points were deducted for either error. At the short and long latency, a target appeared on the screen in the correct location associated with short and long latencies, respectively. If the subject caught the target, the color of the target was green; if the subject missed the target, the color was red. Because of this feedback, the subjects re-experienced the short and long durations every trial. Each subject participated in at most 20 sessions and each session had 3 phases.

**Phase 1: Training.** The + appeared in the middle of the screen and stayed on for either the short or long latency. On the first 2 presentations, the subjects were told the correct answer (the screen printed either “Short” or “Long”). For the next 10 trials, the subjects had to classify the duration of the + by pressing either “S” or “L” after its termination. The subjects were given feedback after each trial. This phase ensured that the subjects learned/re-experienced the reference durations before their testing trials.

**Phase 2: Practice.** Subjects were asked to participate in the experiment as previously described. No points were rewarded or deducted. The probability of the short trial [ $p(S)$ ] of this practice phase matched the probability in the testing phase. In this phase, the subjects were expected to form an estimate of the variability in their representation of elapsed time—the unlikely event that they did not bring an estimate with them to the experiment.

**Phase 3: Testing.** Testing consisted of 500 trials. Although subjects were allowed to take a break after every 10 trials, they were prompted to take a break after every 100 trials. The testing phase was identical to the practice phase except that the goal of this phase was the maximization of the subject’s score. The score was updated after every trial, and the results were shown to the subject according to their assigned type-of-feedback condition. For the first type, numerals represented their score. In this condition, after a given trial duration (short or long), the target that appeared on the side associated with that trial type was a filled square. In the other condition, the target was a set of filled circles, and the number of circles represented how many points had been gained or lost. The subjects in this feedback condition were shown their updated score by the length of a line that ran across the screen. The cumulative score got longer after the subject was rewarded and shorter after the subject was penalized. The color of the line was green, if the cumulative score was above zero, and red otherwise. Payoffs (rewards/penalties) ranged from  $-50$  to  $50$  points. The probability of a short trial was also manipulated. Both of these manipulations were made only between sessions. Type of feedback had no effect.

**Mice.** This procedure is an adaptation of the switch paradigm described by Balci *et al.* (24). The illumination of the control hopper (opposite the 2 feeding hoppers) signaled that a trial could be initiated. A mouse initiated a trial by a poke into this illuminated control hopper. This response requirement ensured that the mouse was in a fixed location at the start of a trial. The trial-initiating poke turned off the light in the control hopper, turned on the lights in both feeding hoppers, and turned on a white noise. The noise went off at the end of the fixed interval for that trial (either 3 s or 6 s). The hopper lights turned off after a nose poke on either hole at or after the termination of the noise. A reward was delivered only if the first



**Fig. 4.** The effects of the relative probability of short and long trials versus the effects of different levels of timing variability (noise, uncertainty) on the expected gain curve. Vertical lines show the modes (locations of the optimal target time,  $T_0$ ). The computations were made with  $g(s < S) = -1$ ,  $g(s > S) = 1$ ,  $g(s < L) = 1$ ,  $g(s > L) = -1$ . The locations of the optimal targets do not change if the negative gains are made 0. The solid curves are computed with  $w = 0.1$ , that is,  $\pm 10\%$  temporal precision; the dashed curves are computed with  $w = 0.2$ , that is,  $\pm 20\%$  temporal precision. The different colors indicate different relative probabilities of short and long trials: Red =  $p(S) = 0.1$ , long trials are 9 $\times$  more likely than short trials; black =  $p(S) = 0.5$ , the two kinds of trials are equally likely; green =  $p(S) = 0.9$ , short trials are 9 $\times$  more likely than long trials.

nose poke at or after the termination of the noise was in the appropriate hopper. After each trial, there was a 30-s fixed delay plus a variable interval of 60 s (drawn from an exponential distribution) before the next trial could be initiated. Although, in theory, mice could wait for the termination of the temporal signal before emitting the short or long latency responses, this did not occur during the experiment.

The feeding hopper associated with the short duration was counterbalanced across mice. Phases 1–4 served as training. Phases 5–7 served as the testing phases and were the phases used in the analysis. Each phase continued until the performance of the mouse stabilized. Unlike the human experiment, the mouse experiment did not manipulate the payoff matrix. Only the probability of a short trial was manipulated across phases. Only the last 10 sessions of the phases were included in the analysis.

**Phase 1.** Three days before the experiment, the mice were food-deprived to 85% of their free-feeding weight. In addition, on each of these 3 days the mice were placed in the experimental chambers for 30 min, and five 20-mg Noyes food pellets (PJAI-0020) were left in their home cages to familiarize them with the pellets used in the experiment.

**Phases 2–4.** We have found that mice do not readily discriminate a 1:2 duration ratio when it is used for the initial training set. Thus, we started training with a 1:3 duration ratio, where the short duration was 3 s and the long duration was 9 s. After training with this pair of temporal intervals, and after several parameter manipulations (i.e., changing the probability of a given trial’s occurrence), the trial duration ratio was reduced to 1:2, using a 3-s short trial and a 6-s long trial. These values were used throughout the rest of the experiment.

**Phase 5.** In phase 5, the probability of a given trial being short was 0.5. This phase was run for 22 daily sessions.

**Phase 6.** In phase 6, the probability of a short trial was 0.25 for half the mice (Group 1) and 0.75 for the other half (Group 2). This phase was run for 20 daily sessions.

**Phase 7.** In phase 7, the probability of a short trial for Group 1 was 0.9, whereas for Group 2 the probability was 0.1. This session was run for 22 daily sessions.

**Optimal Decision-Making Model.** The data were compared to a model in which animals switch from the short location to the long location at the time that maximizes their expected gain. The expected gain function is

$$E(g) = g(s < S)p(S)\Phi(S, T, wT)$$

