

Decision From Models: Generalizing Probability Information to Novel Tasks

Hang Zhang, Jacienta T. Paily, and Laurence T. Maloney
New York University

We investigate a new type of decision under risk where—to succeed—participants must generalize their experience in one set of tasks to a novel set of tasks. We asked participants to trade distance for reward in a virtual minefield where each successive step incurred the same fixed probability of failure (referred to as *hazard*). With constant hazard, the probability of success (the survival function) decreases exponentially with path length. On each trial, participants chose between a shorter path with smaller reward and a longer (more dangerous) path with larger reward. They received feedback in 160 training trials: encountering a mine along their chosen path resulted in zero reward and successful completion of the path led to the reward associated with the path chosen. They then completed 600 no-feedback test trials with novel combinations of path length and rewards. To maximize expected gain, participants had to learn the correct exponential model in training and generalize it to the test conditions. We compared how participants discounted reward with increasing path length to the predictions of 9 choice models including the correct exponential model. The choices of a majority of the participants were best accounted for by a model of the correct exponential form although with marked overestimation of the hazard rate. The decision-from-models paradigm differs from experience-based decision paradigms such as decision-from-sampling in the importance assigned to generalizing experience-based information to novel tasks. The task itself is representative of everyday tasks involving repeated decisions in stochastically invariant environments.

Keywords: decision under risk, decision from experience, constant hazard rate, exponential survival function, generalization

Supplemental materials: <http://dx.doi.org/10.1037/dec0000022.supp>

Decision-making is often modeled as a choice among lotteries. A lottery is a list of mutually exclusive possible outcomes, O_i , $i = 1, \dots, n$, with corresponding probabilities of occurrence p_i , $i = 1, \dots, n$, such that $\sum_{i=1}^n p_i = 1$. A decision maker may, for exam-

ple, be constrained to choose one of the following:

Lottery A: 50% chance to win \$900 and 50% chance to get nothing; versus

Lottery B: \$450 for sure.

Decision tasks can be divided into classes based on the source of probability information available to the decision maker (Barron & Erev, 2003; Hertwig, Barron, Weber, & Erev, 2004). In the traditional “decision from description” task—illustrated above—(e.g., Kahneman & Tversky, 1979), participants choose between lotteries where the probability of each outcome is explicitly given.

In a new and growing research area known as “decision from experience” or “decision from sampling” (Barron & Erev, 2003; Hadar & Fox, 2009; Hertwig et al., 2004; Ungemach, Chater, & Stewart, 2009; see Rakow & Newell, 2010 for a review), participants learn

This article was published Online First November 3, 2014.

Hang Zhang, Department of Psychology and Center for Neural Science, New York University; Jacienta T. Paily, Department of Psychology, New York University; Laurence T. Maloney, Department of Psychology and Center for Neural Science, New York University.

This research was supported by NIH EY019889 to Laurence T. Maloney.

Correspondence concerning this article should be addressed to Hang Zhang, Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, 2nd Floor, New York, NY 10003. E-mail: hang.zhang@nyu.edu

probabilities by repeatedly sampling outcomes from multiple lotteries and then decide which they prefer.¹ The participant may, for example, repeatedly execute Lottery A above and receive a stream of outcome values \$900 and \$0, with roughly as many of the former as the latter. Sampling from Lottery B leads only to outcomes \$450.

With large enough samples, the information about probability in decision from sampling converges to that available in decision from description.² However, participants in decision from description and decision from sampling experiments seem to treat probabilities very differently. In decision from description, human choices deviate from those that maximize expected utility as if small probabilities were overweighted and large probabilities underweighted (Tversky & Kahneman, 1992). In decision from sampling, the reverse pattern—small probabilities underweighted and large probabilities overweighted—is found (Erev et al., 2010; Hertwig et al., 2004; Teodorescu & Erev, 2014).

Decision from sampling captures many real-world situations where explicit probability information is unavailable but individuals may “try out” the various choices repeatedly before committing to a decision. There are, however, situations where people must make decisions, probabilities are not given, and repeated sampling is either not possible or undesirable. They might not, for example, want to learn how to cross a wide, busy street by trial and error. However, if people can somehow generalize their experience with narrow, safe streets by modeling how their probability of survival varies with traffic density and speed, they are potentially able to estimate probabilities of survival in novel situations (e.g., high speed traffic) without direct experience. Generalization extends the benefits of learning to a wider range of potential decision tasks.

In decision-from-sampling tasks (Barron & Erev, 2003; Hertwig et al., 2004), there is—by design—little basis for generalizing to novel conditions (lotteries). The focus of research is on experience—not generalization. If, for example, the individual samples Lottery A’s outcomes from a red deck of cards and Lottery B’s outcomes from a green deck he has information concerning the frequencies of possible rewards from each of the two decks. He has no basis for

generalizing to an unsampled blue deck of cards representing Lottery C.

In the present study, we introduce a new class of decision task, “decision from models,” where—in order to do well—people must generalize probability information from past experience to a novel task based on an internal model of the probabilistic environment. In our study, on each trial, participants had to choose between two paths through a virtual minefield (Figure 1). A fixed number of invisible mines were randomly and uniformly distributed in the minefield. The probability of successfully traversing a linear path in the field without running into a mine is a decreasing exponential function of path length x (the *survival function*)³:

$$P(x) = e^{-\lambda x}, x > 0 \quad (1)$$

The parameter $\lambda > 0$ is the *hazard rate*. On most trials, the participant had to choose between a shorter, thus safer, path toward a smaller reward and a longer, more dangerous path toward a larger reward.⁴ Surviving the path they chose would result in the specified reward; otherwise, they would receive nothing. The participant could potentially estimate the hazard rate λ from feedback in the training phase. Whether he adopts the correct, exponential model of Eq. 1, though, is a separate question, one we return to below.

The tasks in the test phase differed only in the path lengths and values of the lotteries presented; no feedback was provided. The new conditions and absence of feedback prevented the participant from using stimulus-response

¹ Later on, we will use the term “decision from sampling” rather than the more general term “decision from experience” to refer to this research area and its tasks, to avoid confusion with other decision tasks that have experience-based components.

² One striking result in the decision from sampling literature is that participants tend to take very small samples even when they are free to sample as much as they wish (Hertwig et al., 2004).

³ If the observer takes discrete steps of a fixed size, his length of survival is a geometric distribution. As the steps become smaller and smaller, the limiting distribution is the exponential which we use in the following analysis.

⁴ On a few “catch trials” participants were asked to choose between a shorter path with larger reward and a longer path with smaller rewards. The catch trials were used to screen out participants as described in the *Method* section.

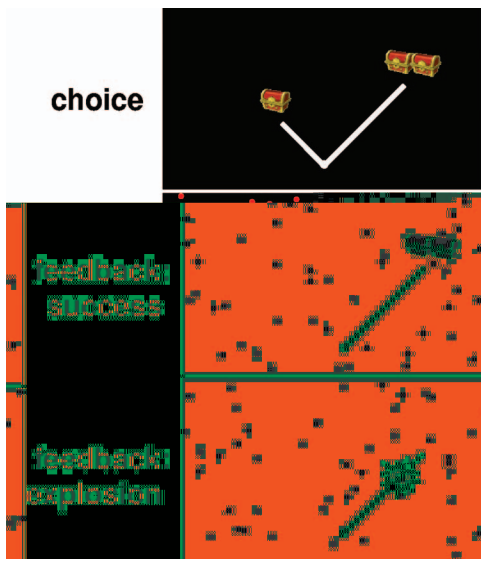


Figure 1. Example of the task. Top: Participants in a virtual minefield chose between a shorter path leading to a smaller reward and a longer path leading to a larger reward. The probability of failure with each successive step—the hazard rate—is constant. Each treasure chest denoted \$1. Middle and bottom: Possible feedback on the chosen path (success and failure). Feedback was present in the training phase and absent in the test phase. See *Method* section.

mappings or “model free” reinforcement learning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) to complete his choices. He had to use the information in the training phase to work out how the probability of survival changes with path length; that is, the model that maps probability information from training to paths of novel lengths in the test phase in Eq. 1. If he could make effective use of Eq. 1 with an accurate estimate λ , then he could precisely predict survival probabilities for any path length.⁵

We test whether the individual acquires and makes use of the exponential survival function of Eq. 1 with correct hazard rate, allowing correct translation of information from the training to the test phase. The internal model, analogous to E. C. Tolman’s cognitive map (Tolman, 1948), would allow people to choose among novel lotteries that they have never had direct experience with.

We do not assume that the internal model would come exclusively from the individual’s experience in training: In the training phase, we

sought to give participants every opportunity to infer the correct model, including instructing them that the mines were randomly distributed and that the probability of surviving a specific path depended only on the path length. They also received feedback that includes display of the full minefield at the end of each training trial. And of course they experienced success or failure in each training trial. Any of this information could have contributed to their learning the correct model.

The task and environment we considered is of interest in itself. It is an example of a *process of constant hazard*: no matter how long an organism has survived, on its next step it has the same chance of encountering a mine as it did on the first. For a process to be a process of constant hazard simply means the environment is *memoryless*: the number of success in past attempts has no influence on the probability of success in the next attempt given that one has survived until now.

Of course, repeated choices in many environments are properly modeled with *changing hazard functions*: repeated successes in the past may increase the probability that the next step will lead to success—or decrease it. Yet—in many environments—repeated choices can be approximated as processes of constant hazard on time or distance or trial. All that is needed is a plausible basis for assuming that each successive step in time or space incurs the same probability of failure: unprotected sex, clicking on an e-mail from an unknown sender, darting across a busy street at lunchtime, moving across a meadow exposed to predators—or crossing a minefield.

People have been found to be surprisingly sensitive to the forms of probability distributions (exponential, Gaussian, Poisson, etc.) they encounter in everyday life: They can accurately estimate the distributions of the social attitudes and behaviors of their group (Nisbett & Kunda, 1985); they can also use appropriate probability distributions for inference or prediction (Griffiths & Tenenbaum, 2006, 2011; Lewandowsky, Griffiths, & Kalish, 2009; Vul, Goodman, Griffiths, & Tenenbaum, 2009). We investigate

⁵ The correct generalizations are somewhat nonintuitive. If the probability of surviving a path of length L is p then the probability of surviving a path of length $2L$ is p^2 .

whether they adopt the exponential survival function (constant hazard) when it is appropriate.

Our focus on model-based transfer also distinguishes the path lottery task from sampling-based decision tasks that have a similar risk structure, such as the Balloon Analogue Risk Task (BART; Lejuez et al., 2002; Pleskac, 2008; Wallsten, Pleskac, & Lejuez, 2005). In the minefield, every tiny step along the path incurred a *fixed* probability of triggering a mine and the participant need to make use of this constraint to infer the probability of survival for novel path lengths in the test phase. In the BART, participants repeatedly pumped a breakable balloon to accumulate reward. They received an amount of reward for each “pump” but would lose all reward if the balloon broke. The decision at any moment was whether to risk one more “pump” for additional reward. Participants were never required to generalize their knowledge of the probability of failure incurred by one “pump” to that of an arbitrary number of “pumps.” Furthermore, in the test phase of the minefield task there was no feedback, precluding any trial-by-trial learning strategies and making the task a rigorous test of model-based transfer.

We analyzed human choices in the path lottery task to answer two questions. First, were people’s choices based on the correct exponential model of probability of survival? If not, what model did they use? Second, do people correctly estimate the parameters of whatever model they assume? That is, if they correctly adopted an exponential survival function, was their estimation of the hazard rate (λ in Eq. 1) accurate? Or, if they assumed a model of a different form, are the parameters of that model appropriately estimated?

Modeling

Modeling Stochastic Choice

Human choices are stochastic (Rieskamp, 2008): People do not necessarily make the same choice when confronted with the same options. To predict the choice behavior in individual trials, we constructed choice models that are stochastic.

Denote the pair of path lotteries on a specific trial as L_1 : (v, x) and L_2 : (w, y) where v, w

denote rewards and x, y denote path lengths. In each choice model, we computed $\Pr(L_2)$, the participant’s probability of choosing L_2 , and modeled the participant’s choice on the trial as a Bernoulli random variable with mean $\Pr(L_2)$ (Erev, Roth, Slonim, & Barron, 2002). For each participant, we fitted the free parameters in each choice model to the participant’s choices using maximum likelihood estimation (Edwards, 1972; see Supplemental Materials).

Except for the resampling choice models (that we present later), $\Pr(L_2)$ is determined by the normalized expected utilities of L_1 and L_2 and a free parameter τ that reflects the randomness in the participant’s choices (Busemeyer & Townsend, 1993; Erev et al., 2002; Sutton & Barto, 1998; see Supplemental Materials).

Utility

The utility of a specific monetary reward was modeled as a power function with parameter $\alpha > 0$ (Luce, 2000):

$$u(v) = v^\alpha \quad (2)$$

Pleskac (2008) has pointed out that the utility function and probability weighting function in the task of Wallsten et al. (2005) were not simultaneously identifiable. We also found that α and the free parameters in some choice models could not be simultaneously estimated (see Supplemental Appendix C online for proof). To overcome this issue, we treated α in all the choice models as a constant for each fit (i.e., not a free parameter in the actual fit) and estimated each choice model under a range of different α ’s that are typical in the decision-making literature (Gonzalez & Wu, 1999; Tversky & Kahneman, 1992). In this way, we could verify that any conclusions we drew were valid across the range of utility functions typically encountered.

v and a longer path of length y toward a larger reward w . For any specific (v, x) and w , we estimate the equivalent length \tilde{y} where the participant is indifferent between the two options, that is, $(v, x) \sim (w, \tilde{y})$ experimentally. We assume an equivalence of expected utilities:

$$u(v)P(x) = u(w)P(\tilde{y}) \quad (3)$$

where $u(\cdot)$ is the utility function and $P(\cdot)$ is the probability of success. Let $\beta = u(w)/u(v)$ ($\beta > 1$ as $w > v$). Equation 3 can then be written as:

$$P(x) = \beta P(\tilde{y}) \quad (4)$$

When the survival function is specified (e.g., Eq. 1), for any specific β , we can compute the equivalent length \tilde{y} , based on Eq. 4, as a function of x . The functional form of \tilde{y} against x is determined by the survival function the participant assumed. Conversely, we can use the measured relationship of \tilde{y} to x to rule out some possible choice models.

We first describe three choice models, which differ in their assumptions concerning the survival function. Their assumptions and predictions are illustrated in Figure 2. We will later consider additional choice models (notably models based on resampling) that are motivated by the observed choice patterns. These will be introduced in the Results section.

Exponential Choice Model

When the hazard rate is constant, the probability of success (survival function) is an exponential function of the path length, as indicated by Eq. 1. The larger the hazard rate λ , the riskier the minefield, and the smaller the probability of success. Substituting Eq. 1 into Eq. 4 and, taking logarithms, we have (see Supplemental Appendix A for derivation):

$$\tilde{y} = x + \frac{\ln \beta}{\lambda} \quad (5)$$

That is, for any specific choice of β , if participants (correctly) assumed the exponential model, their measured \tilde{y} would be a linear function of x whose slope equals 1 (Figure 2, left bottom). Intuitively, if you are indifferent between traveling 1 cm for \$1 and 3 cm for \$2,

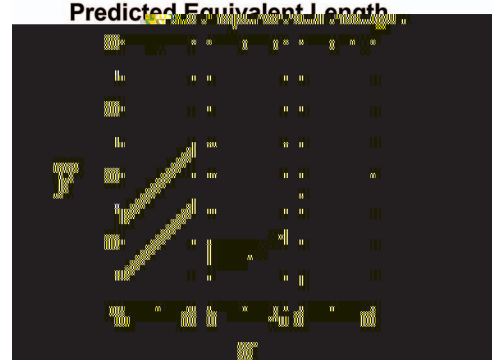
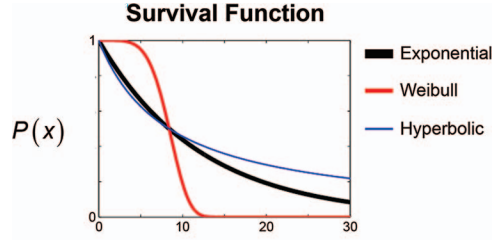


Figure 2. Illustration of three classes of survival models. Top: Survival functions (probability of success as a function of path length) assumed by the models. Black (thickest) line denotes an exponential model. Red (second thickest) line denotes a Weibull model ($\gamma > 1$). Blue (least thick) line denotes a hyperbolic model. Bottom: Predicted equivalent length as a function of the shorter length in the lottery pair. The left, central, and right panels denote the predictions of the exponential, Weibull, and hyperbolic models. Lighter and darker colors in each panel correspond to a higher and a lower β 's. See Eq. 1–9. The unit of length is arbitrary.

you should be indifferent between traveling 11 cm for \$1 and 13 cm for \$2.

Weibull Choice Model

The Weibull survival function (Weibull, 1951) is a generalization of the exponential that includes smoothly increasing and decreasing hazard functions⁶:

$$P(x) = e^{-(\lambda x)^\gamma} \quad (6)$$

⁶ If the participant correctly assumes an exponential survival function but has a probability distortion in the one-parameter Prelec form (Prelec, 1998), his survival function is of the form of a Weibull function. The survival function is probability versus distance and any distortion of the survival function can be treated as a form of probability distortion.

When $\gamma = 1$, it coincides with the exponential. The corresponding Weibull hazard function is

$$h(x) = \gamma\lambda(\lambda x)^{\gamma-1}. \quad (7)$$

which is an increasing power function when $\gamma > 1$, a decreasing power function when $0 < \gamma < 1$, and constant when $\gamma = 1$.

The Weibull equivalent length (see Supplemental Appendix A for derivation) is:

$$\tilde{y}^\gamma = x^\gamma + \frac{\ln\beta}{\lambda^\gamma} \quad (8)$$

For any specific β , it is a curve that converges to the identity line when x approaches infinity (Figure 2, central bottom). The curve is convex when $\gamma > 1$ and concave when $0 < \gamma < 1$.

Hyperbolic Choice Model

There is an evident analogy between our task of trading-off distance for reward and temporal discounting (Frederick, Loewenstein, & O'Donoghue, 2002). The counterpart of constant hazard rate in temporal discounting is constant discount rate. Human decisions on delayed rewards, however, exhibit a declining discount rate (e.g., Thaler, 1981), which is well fit by a hyperbolic model (see Frederick et al., 2002, for a review; Kirby & Maraković, 1995; Myerson & Green, 1995; Raineri & Rachlin, 1993).

Analogously, participants may assume a survival function that declines with distance following a hyperbolic function:

$$P(x) = \frac{1}{1 + \delta x} \quad (9)$$

where $\delta > 0$ is a parameter of discounting rate, analogous to the hazard rate λ in the exponential model. The larger the δ , the smaller the probability of success. Substituting Eq. 9 into Eq. 4, we get a linear relationship (see Supplemental Appendix A for derivation):

$$\tilde{y} = \beta x + \frac{\beta - 1}{\delta} \quad (10)$$

Recall that $\beta = u(w)/u(v) > 1$. That is, the hyperbolic model predicts a line of a slope

greater than one (Figure 2, right bottom). If you are indifferent between traveling 1 cm for \$1 and traveling 3 cm for \$2, you would prefer traveling 13 cm for \$2 to traveling 11 cm for \$1.

Method

Apparatus and Stimuli

Stimuli were presented on a 32-in. (69.8 × 39.2 cm) Elo touch screen, using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Figure 1 shows an example of stimuli. Two paths branched toward the upper left and right of the screen. Each treasure chest at the end of a path represented a reward of US\$1 for completing the path. On each trial, 369 invisible circular mines were independently and uniformly distributed at random. For every 1 cm traveled along a path, there was an approximately 8% probability of hitting a mine.

Procedure and Design

The experiment consisted of two phases: training and test. It lasted approximately 75 min.

Training. The purpose of the training phase was to provide participants an opportunity to familiarize themselves with the task and estimate the hazard implicit in moving through the minefield. On each trial, participants chose between two paths of different rewards and different lengths: (v , x) and (w , y). They were informed that mines were randomly distributed and the probability of surviving a specific path depended only on the path length.

After choosing, the participant was shown an animation of her travel along the chosen path. If she hit a mine, the animation stopped and a graphical explosion marked the location of the mine (Figure 1, bottom). Otherwise the animation continued to the end and the treasure box(es) flashed. Thereafter all the mines were briefly displayed as an aid to the participant in learning the hazard of traveling in the minefield and the correct form of the hazard function.

During training, the rewards (v vs. w) were \$1 versus \$2. On each trial, $P(x)$ was randomly chosen to be between 0.3 and 0.7, and $P(y)$ was either 3/8 of $P(x)$ or 2/3 of $P(x)$ with equal likelihood. Accordingly, x was between 4.3 and 14.5 cm and y was 11.8 or 4.9 cm longer than x .

There were 160 training trials (including 20 screening trials as described below). Participants were not told that there would be a subsequent test phase with no feedback.

Test. In the test phase, the task was the same except that feedback was withdrawn and the values of the rewards changed. There were two reward conditions (v vs. w): \$1 versus \$3, and \$2 versus \$3. The x value could be 11.1, 8.4, 6.2, 4.3, or 2.7 cm so that $P(x)$ equaled 0.4, 0.5, 0.6, 0.7, or 0.8. For each of the these 10 conditions, y was adjusted by a 1-up/1-down adaptive staircase procedure of multiplicative steps and we computed the equivalent length \tilde{y} as the geometric mean of the x after the first two reversals. All staircases were interleaved and terminated after 60 trials, resulting in $10 \times 60 = 600$ test trials.

It is known that people accurately estimate length for lines ranging from 1 cm to 1m (Teghtsoonian, 1965). The path lengths we chose to use were well within this range, so there would be little chance that biases would arise from a misperception of path lengths.

Reward. After each phase, a few trials (one out of training, six out of test) were chosen at random and the participant's was rewarded for the outcome of traveling along the chosen path. The sum of winnings on these reward trials was paid to the participant as a bonus. The participant knew, at the beginning of training, that some of the trials would be rewarded in this way. The participant received US\$12 per hour in addition to the bonus.

Screening. In 20 trials of the training phase, the shorter path led to a *larger* reward and the longer path to the *smaller* reward. A participant who selected the higher risk path to the smaller reward has violated dominance. Any participant who failed to choose the nondominant option in more than 4 of the 20 trials was excluded from the remainder of the experiment.

Participants

Data from 17 participants—4 men and 13 women, aged 18–50, median = 21—were analyzed. An additional two participants failed the dominance screening and did not complete the experiment. We determined the sample size in advance. No variables or conditions were dropped. The experiment was approved by the

University Committee on Activities Involving Human Subjects at New York University.

Results

Choices in the Training

In the training phase of the experiment, participants chose between pairs of path lotteries $L_1: (v, x)$ and $L_2: (w, y)$. We use L_2 to denote the lottery that had the larger reward and the longer path length. The larger reward was always twice the smaller reward (\$2 vs. \$1). The probability $P(y)$ of surviving the longer path was $2/3$ of the probability $P(x)$ of surviving the shorter path for 70 trials and $3/8$ of the probability $P(x)$ of surviving the shorter path for an additional 70 trials, all trials interleaved. Figure 3A shows how the participants' mean probability of choosing L_2 evolved with experience. In both conditions, the participants showed a trend toward picking the less risky option L_1 . This trend, however, was not significant, according to a 2×5 repeated-measures ANOVA, $F(4, 144) = 1.57, p = .19$.⁷

We did not model the choices in training. What we can infer from Figure 3A, without any further assumptions on the underlying decision process, is that there was no significant indication that the participants were becoming more and more risk-seeking during their training. We will return to this point in the Discussion.

Measured Equivalent Lengths in the Test

Before testing participants' choice behavior against the choice models we introduced earlier, we would like to give an overview of their choice patterns. In the test phase, we measured the equivalent length \tilde{y} , the length for which the participant was indifferent between the two options, that is, $(v, x) \sim (w, \tilde{y})$. The mean equivalent length across participants, plotted as dots in Figure 3B, visually agrees with the pattern predicted by a survival model of the correct exponential form (Eq. 4): \tilde{y} is a linear function of x whose slope is one.

The lines in Figure 3B shows the predictions of the correct model, an exponential model whose hazard rate was estimated from the ran-

⁷ The degrees of freedom correspond to $2 P(y)/P(x)$ conditions by 5, the number of bins of 14 trials each.

dom samples each participant actually saw using the maximum likelihood method. The measured equivalent lengths were smaller than those predicted by the correct model. Participants appeared to be highly risk-averse: When choosing between a shorter path leading to \$1 and a longer path leading to \$3, at the indiffer-

Table 1
Model Comparison: Constant Versus Changing Hazard Rate

Model	<i>df</i>	Assuming $u(v) = v$		Assuming $u(v) = v^{0.88}$		Assuming $u(v) = v^{0.49}$	
		No. of best fits	Mean BIC	No. of best fits	Mean BIC	No. of best fits	Mean BIC
Exponential	2	12	-404.7	11	-404.7	11	-405.2
Weibull	3	4	-404.9	4	-404.9	4	-405.4
Hyperbolic	2	1	-418.3	2	-417.9	2	-414.3

Note. BIC = Bayesian information criterion. *df* refers to the number of free parameters in the model. *No. of best fits* refers to the number of participants that were best fit by the model.

utility function. Among the three choice models, most participants were best fit by an exponential model (constant hazard rate); the runner-up was the Weibull model; few participants were best fit by the hyperbolic model.

The Weibull model is a generalized form of the exponential model, which allows the hazard rate to be increasing ($\gamma > 1$), constant ($\gamma = 1$), or decreasing ($\gamma < 1$) over distance. We used the estimated γ in the Weibull model to provide additional information about the hazard-changing trend assumed by the participant. Figure 4A shows γ for each participant, assuming $u(v) = v$: The γ of most participants was close to one (median = 1.16), implying an assumption of constant hazard. Across participants there was no trend away from constant hazard.

Estimation of hazard. In the training phase, participants observed whether and where their chosen paths hit mines. Based on the random samples that each participant actually saw, we estimated the hazard rate λ_0 that most likely generated these samples using the maximum likelihood method. It ranged from 0.069 to 0.091 (median = 0.078); if the participant's choices were consistent with this hazard rate we would judge that she correctly estimated hazard during training.

Figure 4B shows the estimated hazard rate λ when the participant's choices were fitted to the exponential model, assuming $u(v) = v$, relative to her true hazard rate λ_0 . Most participants overestimated the hazard rate. The median ratio was 3.53, 3.13, and 2.41, respectively, for $u(v) = v$, $u(v) = v^{0.88}$, and $u(v) = v^{0.49}$.

Learning-Based Choice Models

What would lead participants to overestimate their hazard rate? We constructed a delta-rule learning algorithm to estimate the survival func-

tion based on experience in the training phase (Online Supplemental Materials). The algorithm could potentially result in the observed overestimation of hazard rate. The basic idea was that positive and negative outcomes could be weighted in an unbalanced way. In particular, the overestimation of hazard rate was implemented as an overweighting of negative outcomes (running into a mine) compared with positive outcomes (survival).

The survival function produced by the learning algorithm is a step function. We call the choice model based on it the *learning-based nonparametric model*. We constructed a second learning-based model, the *learning-based exponential model*, whose survival function is simply a smoothing of the survival function of the nonparametric model. Remarkably, most participants were better fit by the learning-based exponential model than by the learning-based nonparametric model (Table 2), providing further evidence that participants assumed survival models of the correct exponential form.

Survival-Based Versus Resampling Choice Models

Instead of building a one-to-one mapping between path length and probability of survival, as implemented in the survival-based choice models discussed earlier, participants might memorize all the path lotteries they attempted during training together with their outcomes and use these instances to estimate the probability of survival for each test path. However, participants could not simply recall paths of the same length as the test path because there was no exact match from the training set—the path lengths in the training set were randomly chosen from a continuum. The issue then is, how can the participant generalize the information

gained in the training phase? This information can be thought of as ordered triples (x, x', O) where, x is the path length on a trial, x' is the path length traveled until a mine was encountered on a failure trial or the total length of the path on a reward trial, and O represents the observed outcome (success or failure).

We explored simple rules based on resampling from the information gathered during training, stored in memory. We constructed three *resampling models*, all of which assume that participants randomly (re)sample from their training experience. The models differ in the memory populations to resample from and the generalization rule applied to the resampled paths.⁸

The first resampling model assumed that participants resample only from paths that are longer than the test path. Each resampled path that ran into a mine before reaching the length of the test path corresponded to a possible failure for the test path (the test path would have encountered the same mine as the resampled path if placed on the same minefield). The probability

of survival for the test path was computed as the number of survivals averaged across all the sampled paths. We call this model the *unbiased resampling model*: the estimated probability of survival for any test path would not deviate from the true probability of survival in any systematic way.

To accommodate possible biases in participants' choices, we considered two additional resampling models that involve biased (but plausible) usage of the resampled paths. Both of them assumed that resamples could be drawn from any path the participant has experienced in training. When the resampled path length was longer than the test path, the rule was the same

⁸ In statistics, resampling is a method based on drawing

Table 2
Model Comparison: Learning-Based Models

Model	df	Assuming $u(v) = v$		Assuming $u(v) = v^{0.88}$		Assuming $u(v) = v^{0.49}$	
		No. of best fits	Mean BIC	No. of best fits	Mean BIC	No. of best fits	Mean BIC
L-B non-parametric	3	4	-411.5	5	-410.8	5	-409.1
L-B exponential	3	13	-407.9	12	-407.9	12	-408.4

Note. L-B abbreviates for learning-based. df refers to the number of free parameters in the model. *No. of best fits* refers to the number of participants that were best fit by the model. BIC = Bayesian information criterion.

as in the unbiased resampling model. When the resampled path was shorter than the test path, the *optimistic resampling model* assumed that the resampled path counted as a success if the resampled path itself ended in success, that is, success on the shorter, resampled path was treated as a guarantee of success on the longer path; conversely, the *pessimistic resampling model*, treated the same case as a failure. The probability of survival is typically overestimated in the optimistic resampling model and underestimated in the pessimistic model.

Could participants' choice behaviors arise from a resampling process? We compared the fits of the survival-based choices models with the resampling models. As shown in Table 3, only 3 out of 17 participants could be better captured by a resampling model than by a survival-based model.

We noticed three facts. First, on average the resampling models fit much worse to participants' choice behaviors than most of the model-based models did (mean BIC in Table 3). Second, the pessimistic resampling model fit better than the other two resampling models (number of best-fit participants). Recall that the pessi-

mistic resampling model would yield an underestimation of the probability of survival, that is, an overestimation of hazard rate. Third, for all the participants who were best fit by a resampling model, the fitted sample size was 1, in contrast to a typical sample size of 5 for sampling-based decisions (Erev et al., 2010).

The BIC difference between the survival-based models and the resampling models was smaller for $u(v) = v^{0.49}$. At first glance it appears that the resampling models do "relatively less poorly" when the utility function is most concave. This is simply because of the fact that larger rewards are associated with longer path lengths (lower probabilities) and failure to generalize correctly is not punished as much if larger rewards are scaled to be relatively less important.

Discussion

In the present article, we introduce decision from models as a class of decisions parallel to decision from description (Tversky & Kahneman, 1992) and decision from sampling (Barron & Erev, 2003). It has an experienced-based

Table 3
Model Comparison: Survival-Based Versus Resampling Models

Model	df	Assuming $u(v) = v$		Assuming $u(v) = v^{0.88}$		Assuming $u(v) = v^{0.49}$	
		No. of best fits	Mean BIC	No. of best fits	Mean BIC	No. of best fits	Mean BIC
Survival-based		14		14		14	
Exponential	2	10	-404.7	9	-404.7	9	-405.2
Weibull	3	3	-404.9	3	-404.9	3	-405.4
Hyperbolic	2	1	-418.3	2	-417.9	2	-414.3
Resampling		3		3		3	
Unbiased	1	0	-434.7	0	-434.8	1	-427.9
Optimistic	1	0	-479.7	0	-479.7	0	-479.5
Pessimistic	1	3	-437.9	3	-437.4	2	-430.6

Note. BIC = Bayesian information criterion. df refers to the number of free parameters in the model. *No. of best fits* refers to the number of participants that were best fit by the model.

component but is focused on generalization: participants need to develop a model allowing them to generalize probability information from their past experience to novel tasks. We investigated whether people could correctly learn an exponential survival model and base their choices in novel tasks on the correct model.

To maximize their expected reward, participants needed to take the information acquired during training and use it to generalize to novel conditions. If their memory for this information is distorted then their choice of model and/or estimate of its hazard rate parameter could be affected. But remembering this information accurately is not sufficient to guarantee that they pick the correct model or correctly use it to generalize.

We found that most participants made choices consistent with a survival model of the correct exponential form: The exponential model fit better to participants' choices than the Weibull model and the hyperbolic model.

Prelec and Loewenstein (1991), and Green and Myerson (2004) suggested a number of parallels between decision under risk and intertemporal decision. The correct exponential functional form provided better fits to participants' data than the hyperbolic form found in the temporal discounting literature. If risk in our task were exchangeable with temporal delay, we would have expected a hyperbolic discounting of distance (i.e., hyperbolic survival function) in the minefield. However, hyperbolic internal models were favored by fewer than one eighth of the participants.

Evidence in favor of the correct assumption of the exponential form also came from a model comparison with a nonparametric model. We "piped" the participant's training experience through a specific learning algorithm and found that, if the resulting survival function was then smoothed to approximate an exponential function (no parameter added), it would provide a better fit for most participants than that of the survival function originally learned. That is, choice models that assumed internal models of the correct exponential form better captured participants' choices than those that assumed no specific functional forms.

We further verified that participants' choices in the test could not be reduced to a resampling process that is analogous to decision from sampling but based on resampling from memory

(Erev et al., 2010). The performance of most participants were better accounted for by choice models that rely on a survival function than by resampling models. Even for the few participants who may have depended on resampling, the sampling process was based on very small samples. The sample size was typically 1, even smaller than the number of samples (~ 5) found in decision-from-sampling tasks (Erev et al., 2010). This result echoes Vul et al.'s (2009) claim that a sample size of one is efficient for inference from probability distributions.

But most participants did not base their choices on the correct model exactly: they overestimated the hazard rate of the exponential survival function. They did not go as far as they should go when offered a larger reward. According to participants' average response in the \$1 versus \$3 condition, the expected gain of the shorter path was only 49% of that of the longer path. We noticed a similar "risk-averse" outcome in the BART: Their participants pumped the balloon fewer times than would maximize expected gain (Lejuez et al., 2002). When the optimal number of pumps was 64, participants on average pumped only 37.6 times (Wallsten et al., 2005).

The BART and our path lottery task, put together, shed light on this overestimation of hazard. Wallsten et al. (2005), in their modeling of the BART, attributed it to an incorrect prior belief that would be corrected by experience. Our results cast doubt on this explanation. Since the hazard rates of both tasks were arbitrarily chosen by the experimenter, why should people have a prior hazard rate in both tasks that is higher than the true? Moreover, if participants in the path lottery task did have an overestimation of hazard before the task but were able to correct it through experience, we would expect an increasing trend in their probability to choose the more risky option (L_2) during training. This was not the case (Figure 3A). We conjecture that participants' overestimation of hazard reflects an improper estimation of probability information from their experience and not simply lack of experience.

It is natural to expect people to correctly estimate both the correct functional form and the correct parameters (as in Griffiths & Tenenbaum, 2006) or neither. Mozer, Pashler, and Homaei (2008) pointed out the possibility that the seemingly accurate human predictions for a

specific distribution may be based on a few memorized samples of the distribution. A similar possibility can hardly be true for our decision task. If we assume that participants' memory of samples was faithful and that they combine sample information correctly, their decisions should not systematically deviate from the predictions of the true hazard rate. Conversely, if their memory of samples were distorted, there would be no reason to assume that the form of the distribution would be unaffected. And as we demonstrated earlier, resampling models did not provide good fits to participants' choices. Of course, we did not test all possible resampling models and could not exclude the possibility that a different generalization model based on resampling such as the similarity-based sampling model of [Lejarraga and Gonzalez \(2011\)](#) might work better than the three sampling models we tested. However, we see no way to develop a sampling-based model that could account for the coexistence of the correct functional form and the overestimation of hazard.

Two basic approaches to learning a specific functional relationship are (1) to rely on a set of hard-wired base functions or (2) to use associative learning free of functional forms ([McDaniel & Busemeyer, 2005](#)). The disassociation of learning of the form and the parameter of the hazard function in our task, however, poses something of a challenge for associative learning for the similar reason we discussed above for the sampling approach. We conjecture instead that the exponential (or an approximation to the exponential) is one of the hard-wired functions in the cognitive "repertoire" of functions permitting extrapolation and interpolation. Selection of a particular function (exponential or otherwise) is then seen to be analogous to contour completion in human perception ([Metzger, 2006](#)). We have argued that processes of constant hazard are appropriate models of many repeated tasks in everyday life, consistent with the claim that the exponential is hard-wired.

We conjecture that participants were able to choose the correct form of model to use simply based on their understanding of the minefield and possibly a few trials of experience. There is evidence that in their use of probability information people take into account how stochastic events or processes are generated physically

([C. S. Green, Benson, Kersten, & Schrater, 2010](#); [Pleskac, 2008](#)). However, people also stick to particular probabilistic models in spite of extensive exposure to contradictory evidence. For example, people incorrectly assume an isotropic model of their motor error distribution after 300 trials of exposure to the vertically elongated true distribution ([Zhang, Daw, & Maloney, 2013](#)). In the BART ([Lejuez et al., 2002](#)), the probability of breaking the balloon with the next pump increased with the number of pumps, but most participants were better modeled as assuming a constant hazard rate ([Wallsten et al., 2005](#)).

In sum, the previous findings allow us to claim decision from models as a new class of decisions that are distinct from decision from description and decision from sampling. People do make choices that generalize to novel tasks using models that allow them to translate experience into accurate probability estimates associated with the novel tasks.

The major goal of the present study was to evaluate decision from models at the computational-theory level in Marr's scheme ([Marr, 1982](#), p. 25). In addition, a process model is needed to account for the patterned choice behaviors we have observed, in particular, how people interpolate and estimate hazard rate as a function of path length and why they sometimes get it wrong.

There are many open questions in this new class of decision tasks. For example, our results do not enable us to ascribe participants' deviations from optimality (maximizing utility) to errors in learning (training phase) or choice (test phase). Further work is needed to elucidate the source of error: what is learned versus how it is used. Another untreated problem is probability distortion. As we noted previously, any distortion of the survival function is equivalent to applying a probability distortion function to the correct, exponential survival function.

Last of all, turning a finite set of observations into a predictive mechanism for an infinite number of possible cases is the problem of induction. An exciting direction for future research would be to compare tasks with different, correct generalization functions (such as minefields with nonuniform distributions of mines) to test human ability to generalize correctly.

References

- Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215–233. <http://dx.doi.org/10.1002/bdm.443>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. <http://dx.doi.org/10.1163/156856897X00357>
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459. <http://dx.doi.org/10.1037/0033-295X.100.3.432>
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*, 1204–1215. <http://dx.doi.org/10.1016/j.neuron.2011.02.027>
- Edwards, A. W. F. (1972). *Likelihood*. New York, NY: Cambridge University Press.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the Bootstrap*. New York: Chapman & Hall.
- Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., . . . Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, *23*, 15–47. <http://dx.doi.org/10.1002/bdm.683>
- Erev, I., Roth, A., Slonim, R., & Barron, G. (2002). *Combining a theoretical prediction with experimental evidence*. Retrieved from http://papers.ssrn.com/abstract_id=1111712
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*, 351–401. <http://dx.doi.org/10.1257/jel.40.2.351>
- Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129–166. <http://dx.doi.org/10.1006/cogp.1998.0710>
- Green, C. S., Benson, C., Kersten, D., & Schrater, P. (2010). Alterations in choice behavior by manipulations of world model.

- 64, 263–276. <http://dx.doi.org/10.1901/jeab.1995.64-263>
- Nisbett, R. E., & Kunda, Z. (1985). Perception of social distributions. *Journal of Personality and Social Psychology*, 48, 297–311. <http://dx.doi.org/10.1037/0022-3514.48.2.297>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. <http://dx.doi.org/10.1163/156856897X00366>
- Pleskac, T. J. (2008). Decision making and learning while taking sequential risks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 167–185. <http://dx.doi.org/10.1037/0278-7393.34.1.167>
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66, 497–527. <http://dx.doi.org/10.2307/2998573>
- Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science*, 37, 770–786. <http://dx.doi.org/10.1287/mnsc.37.7.770>
- Raineri, A., & Rachlin, H. (1993). The effect of temporal constraints on the value of money and other commodities. *Journal of Behavioral Decision Making*, 6, 77–94. <http://dx.doi.org/10.1002/bdm.3960060202>
- Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, 23, 1–14. <http://dx.doi.org/10.1002/bdm.681>
- Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1446–1465. <http://dx.doi.org/10.1037/a0013646>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464. <http://dx.doi.org/10.1214/aos/1176344136>
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Teghtsoonian, M. (1965). The judgment of size. *The American Journal of Psychology*, 78, 392–402. <http://dx.doi.org/10.2307/1420573>
- Teodorescu, K., & Erev, I. (2014). On the decision to explore new alternatives: The coexistence of under- and over-exploration. *Journal of Behavioral Decision Making*, 27, 109–123. <http://dx.doi.org/10.1002/bdm.1785>
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8, 201–207. [http://dx.doi.org/10.1016/0165-1765\(81\)90067-7](http://dx.doi.org/10.1016/0165-1765(81)90067-7)
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208. <http://dx.doi.org/10.1037/h0061626>
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323. <http://dx.doi.org/10.1007/BF00122574>
- Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, 20, 473–479. <http://dx.doi.org/10.1111/j.1467-9280.2009.02319.x>
- Vul, E., Goodman, N. D., Griffiths, T. L., & Tenenbaum, J. B. (2009, July). *One and done? Optimal decisions from very few samples*. Paper presented at the Proceedings of 31st Annual Meeting of the Cognitive Science Society, Amsterdam, Netherlands.
- Wallsten, T. S., Pleskac, T. J., & Lejuez, C. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychological Review*, 112, 862–880. <http://dx.doi.org/10.1037/0033-295X.112.4.862>
- Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18, 293–297.
- Zhang, H., Daw, N. D., & Maloney, L. T. (2013). Testing whether humans have an accurate model of their own motor uncertainty in a speeded reaching task. *PLoS Computational Biology*, 9, e1003080. <http://dx.doi.org/10.1371/journal.pcbi.1003080>

Received April 30, 2013

Revision received July 27, 2014

Accepted August 20, 2014 ■