
An analysis of decision under risk in mice, rats and humans

Xiaoyue Zhu^{1,2}, Joshua Möller-Mara^{1,2}, Evgeniya Lukinova^{1,2}, Sylvain Dubroqua^{1,2}, Jeffrey C. Erlich^{1,2,3,*}

1 NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, China

2 NYU Shanghai, Shanghai, China

3 Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), East China Normal University, Shanghai, China

* jerlich@nyu.edu

Abstract

We developed a nonverbal task that allowed us to decompose distinct elements of risk-tolerance in mice, rats and humans. Our goal was to quantify the similarities and differences across species as a foundation for using mice and rats as model organisms for future investigations of decisions under risk. On each trial of the task, subjects decided whether to play a lottery or to get a small certain reward. The lottery magnitude varied independently across trials, and was indicated to the subjects by an auditory cue. We quantitatively explored the strategies of subjects using many functional forms of risk attitude, including mixture models which combined standard economic and financial conceptions of risk with heuristic forms of risk. While data from human subjects were well-described by the majority of the models, rodent behavior was better fit by a model that also includes stimulus-independent biases and trial-history parameters. Once we isolated the history-dependence and stimulus-independent biases, the distribution of utility functions and variance aversion across species became more overlapped. These results provide support for using mice and rats to examine the neurobiology of risk-tolerance, while caution against the use of overly simplified models which ignore certain behavioral tendencies in rodents.

Introduction

Decision making under risk is a topic that attracts wide cross-disciplinary interest from economics, finance, psychology, ecology and neuroscience. The everyday usage of the term ‘risk’ often invokes a perception of potential loss. According to the Oxford English Dictionary, risk is ‘the possibility of something bad happening at some time in the future; a situation that could be dangerous or have a bad result’. This differs from how the term is used by economists, who defined risk as a quantity with known outcome values and *known* outcome probabilities, contrasting ‘ambiguity’, which refers to a quantity with known outcome values and *unknown* probabilities (Knight, 1921). Thus, the economic definition of risk concerns with behavior under known probabilities and entails that the possible outcomes can be exclusively positive.

Understanding how individuals make decisions under risk is of substantial interest from a public health and welfare perspective: excessive risk-taking is associated with drug and gambling addiction (Ahmed, 2018), dangerous teen driving (Williams, 2003) and other impulsive behavior, such as binge-eating and substance abuse (Clifton et al., 2018). On the other hand, inadequate risk-taking is not so desirable either: people who avoid investing in the stock market can have their savings diminished by inflation; a mouse that is unwilling to risk predation for foraging will starve. Data from twin and genome-wide association studies (Xuan et al., 2017; Rao et al., 2018; Anokhin et al., 2009) suggest that genetics accounts for a moderate proportion (~ 30%) of variation in risk-taking. Human choices under risk derive, at least in part, from the same mechanisms evolved in other animals in response to the stochasticity of their natural environment. As such, a robust animal model can help establish the link between genes, brains and risky choice behavior.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

Risk attitude determines whether risk enhances or reduces the value of option for an individual (Tobler and Weber, 2013). To seek a quantitative definition of it, we need to look within different models of risky choice. In general, models of risky choice frame choice options either as outcome-probability pairs, such as in expected utility theory (Von Neumann and Morgenstern, 1953) or outcome distributions, such as in the mean-variance models (Markowitz, 1968). Expected utility theory posits that a rational decision maker should choose in accordance with the theorem of utility-maximization. Utility, a pivotal concept in economics, refers to the subjective satisfaction after receiving or consuming a good or service. When faced with two gamble options, a utility-maximizer first computes the expected utility (EU) for each option, the average utility of different outcomes weighted by their probabilities, and then chooses the one with higher EU. In this framework, the shape of the utility function determines the risk attitude. When the utility function is linear, the decision maker is risk-neutral. She would be indifferent between receiving a guaranteed x (surebet) and receiving $2x$ or nothing with equal probabilities. When the utility function is concave, the decision maker is considered risk-averse and risk-seeking if her utility function is convex. On the other hand, the mean-variance models view risk as the degree of uncertainty inherent in known outcome distributions. Thus, risk as variance is highest when the outcome is most uncertain and decreases as the same outcome becomes increasingly certain. This form of risk is widely used in financial portfolio selection, where the objective is to strike a balance between the maximization of profit (expected value) and minimization of risk (outcome variance) (Markowitz, 1968; Sharpe, 1964). Here, risk attitude is synonymous with one's attitude towards lottery variance. For example, a variance-averse (risk-averse) individual would prefer 50/50 chance of receiving \$20/\$40 over a 50/50 chance of receiving \$10/\$50, even though options have the same expected value ($(20 + 40)/2 = (10 + 50)/2 = 30$).

Using these definitions of risk attitude, studies have found that humans are generally risk-averse in the domain of gains (Kahneman and Tversky, 1979; Holt et al., 2002; Haushofer and Fehr, 2014), although considerable variation among populations of different occupations (Hartog et al., 2002; Hill et al., 2019), gender (Levin et al., 1988; Eckel and Grossman, 2002), and age groups (Tymula et al., 2013; Rutledge et al., 2016) has been reported. In contrast, studies using non-human primates have reliably found them to be risk-seeking (Hayden and Platt, 2007; Heilbronner et al., 2011; O'Neill and Schultz, 2010; So and Stuphorn, 2010) but see Eisenreich and Hayden (2018) for one exception. This risk-seeking behavior has been attributed to the small stakes involved and the repeated occurrences of gambles (Heilbronner and Hayden, 2013). Recently, there is an emerging interest in using rodents to study decision making. Several groups found that rats showed mixed preferences for the safe or risky (variable outcome) option under incomplete (Larkin et al., 2016; Stopper and Floresco, 2011) and cued information (Sugam et al., 2014; Floresco et al., 2018).

Despite concerted effort using human, monkey and rodent subjects to study choice under risk, differences in task design obstruct direct comparison of the 'risk attitudes' measured from these studies. The first difference underlies the so-called description-experience gap: differences between behavior when offers are presented explicitly in writing compared to when they are learned. There is some consensus that that subjects are less risk-seeking when the options are fully described in text, compared to when their outcomes and probabilities are sampled from experience (Hertwig and Erev, 2009; Hertwig et al., 2004; Ludvig and Spetch, 2011). Interestingly, the description-experience gap has been replicated in monkeys (Heilbronner and Hayden, 2016), suggesting that uniquely human cognitive processes such as language do not play a part. The majority of human studies on risky choice used written descriptions, whereas due to their nonverbal nature, animal studies required active sampling of the cue-magnitude / -probability mapping by the animals. Relying on feedback-based reinforcement learning, evidence suggest that individual differences in loss-aversion and learning rates can lead to apparent differences in risk attitudes (e.g. Niv et al., 2012; van Holstein and Floresco, 2020; Zalocusky et al., 2016). The second of such difference highlights a well-known distinction between 'expected uncertainty', referring to known unpredictability in the environment, and 'unexpected uncertainty', the idea that statistics underlying the environment are volatile and prone to change (Dayan and Yu, 2003; Payzan-LeNestour et al., 2013; Soltani and Izquierdo, 2019). Using the probability discounting task where the lottery probability changes without explicit cues, most rodent experiments in fact focused on decision-making under unexpected uncertainty.

We are aware of only a few studies where rats made well-informed choices between a safe and risky option on a trial-by-trial basis (Constantinople et al., 2019b,a). By varying lottery magnitudes and probabilities, Constantinople et al. (2019b) were able to construct rats' psychometric curves and found them exhibit several

behavioral signatures predicted by prospect theory (Kahneman and Tversky, 1979).

Finally, risk attitude is not monolithic. It can be decomposed into elements corresponding to distinct cognitive constructs (Yates, 1992). Risk attitude may well be approximated by the curvature of utility function as in expected utility theory, or variance aversion as in the mean-variance model. However, both models prescribe that a decision-maker fully integrates information through a series of summing and weighting. Is this realistic? As the number of options increases, so does the computational resource required to perfectly integrate them. In fact, many experiments support the view that the human mind relies on simple rule-based strategies, ‘heuristics’, when making decisions under complex information (Kahneman and Frederick, 2002; Brandstätter et al., 2006; Pachur et al., 2013; Venkatraman et al., 2014). The employment of choice heuristics has also been found in monkeys (Kralik et al., 2012), honeybees (Shafir et al., 2002), and starlings (Marsh and Kacelnik, 2002), suggesting a shared mechanism. In many perceptual decision-making studies, rodents exhibited a constant rate of errors independent of the evidence strength known as ‘lapses’ (e.g. Erlich et al., 2015; Carandini and Churchland, 2013; Nikbakht et al., 2018). It was proposed that lapses may reflect a strategic trade-off between exploitation and exploration under uncertainty (Pisupati et al., 2021). Lapse in risky choice manifest as a stimulus-independent bias towards either the surebet or lottery. Thus, previous literature suggest that when modeling risk attitude in animal and human subjects, the researcher needs to consider both lapse and the possible adoption of heuristic strategies.

As discussed above, a common behavioral paradigm is needed to make measured risk attitude across humans and non-human animals comparable. We collected and analyzed trials from mice, rats and humans in a nonverbal task, where subjects made choices guided by auditory cues on a trial-by-trial basis. Although there is substantial literature comparing different functional forms of decision under risk (Heilbronner, 2017; Farashahi et al., 2019; Spitmaan et al., 2019), we are unaware of any previous studies that systematically constructed and compared models that simultaneously estimate all these parameters. Here, we set out to quantitatively explore the strategies of subjects using many functional forms of risk attitude, including mixture models which combined standard economic and financial conceptions of risk with heuristic forms of risk.

Materials and Methods

Animal behavior

Subjects In total, 39 rats and 64 mice were used in this study (Vital River, Beijing, China). Of the 39 rats, 20 were male Sprague Dawley, 19 were Brown Norway (15 males, 4 females). Of the 64 mice, 49 were C57B6 strain (47 males, 2 females) and 15 were CD1 strain (13 males, 2 females). The animals were placed on a controlled-water schedule and had access to free water 20 minutes each day in addition to the water they earned in the task. They were kept on a reversed 12 hour light–dark cycle and were trained during their dark cycle. Animal use procedures were approved by New York University Shanghai International Animal Care and Use Committee following both US and Chinese regulations.

Behavioral apparatus Animal training took place in custom behavioral chambers, located inside sound- and light-attenuated boxes. Each chamber (23 x 23 x 23 cm for rat, 4 x 4 x 4 cm for mouse) was fitted with 8 nose ports arranged in four rows (FIGURE 1A), with a pair of speakers on the left and right side. Each nose port contained a pair of blue and yellow light emitting diodes (LED) for delivering visual stimuli, as well as an infrared LED and infrared phototransistor for detecting animals’ interactions with the port. The port in the bottom row contained a stainless steel tube for delivering water rewards. Animals were loaded and unloaded from the behavioral chambers by technicians daily on a fixed schedule. Each training session lasted for 90 minutes.

Animal risky choice task Trials began with both yellow and blue light-emitting diodes (LED) turning on in the center port. This cued the animal to poke its nose into the center port and hold it there for 1s, after which the center lights were turned off and the choice ports became illuminated. We refer to this period as the ‘soft fixation’ period, because trials did not terminate if the animal withdrew during this period, but had to re-poke into the center poke in order to complete the fixation period. However, during fixation, if the animal

poked into a different port other than the center port, a short white noise would play to indicate that this is a mistake, and we excluded these trials from analyses. After fixation, the choice ports were illuminated.

During the soft fixation period a tone played from both speakers, indicating the lottery magnitude of this trial. We used two different sets of stimuli to indicate the lotteries. In some subjects, the stimuli were pure tones. In other subjects, the stimuli were click trains. There were 5 to 6 distinct frequencies indicating different lottery magnitudes (2.5kHz - 20kHz, 75 dB). Some animals had a positive frequency-to-magnitude mapping, and others had a negative one (some stats here).

The frequency of each pure tone lottery was around one octave away from the adjacent tones, to make distinguishing the different offers perceptually easy Dent et al. (2018). The click trains were also perceptually easy to distinguish. The side of the surebet and lottery port were counterbalanced across the animals. At the end of fixation, the lottery port and surebet port were illuminated with yellow and blue lights, respectively. The tone stopped as soon as the animal made a choice by poking into one of the choice ports. If the animal chose surebet, a small and guaranteed reward would be delivered at the reward port. If the animal chose lottery, it would either receive the corresponding lottery magnitude or nothing based on the lottery probability, which was titrated for an animal and ranged from 0.5 to 0.6 across all subjects. We refer to these trials as 'choice' trials. In order to ensure that the subjects experienced all the outcomes, the choice trials were randomly interleaved with trials that we refer to as 'forced' trials. The forced trials differ from choice trials in that only one of the two ports was illuminated and available for poking, forcing the animal to make that response. The forced surebet and forced lottery trials together accounted for 25% of the total trials. The inter-trial intervals (ITI) were between 3 and 10 seconds. A trial was considered a violation if the animal failed to poke into central 300s after trial start, or it did not make a choice 30s after fixation. Violations were excluded from all analyses, except where they are specifically mentioned. Due to differences in body weight, the rats were given a higher base reward than the mice. On average, a mouse obtained 1.96 ± 0.85 mL rewards in a session, whereas a rat obtained 6.7 ± 2.6 mL.

Training pipeline Animal training took place in two distinct phases: the operant conditioning phase and the risky choice phase. In the operant conditioning phase, naive rats became familiar with the training apparatus and learned to poke into the reward port when illuminated. Trials began with the illumination of reward port, and water reward was immediately delivered upon port entry. After the rats learned to poke in the reward port reliably, they proceeded to the next training stage where they had to first poke into an illuminated choice port (left or right, chosen randomly) before the reward port was illuminated for reward. They graduated to the risky choice phase if they correctly performed these trials at least 40% of the session.

In the risky choice phase, rats started with only two frequencies: the lowest and highest, corresponding to the smallest and largest lottery magnitude (or the other way around for the animals with reversed mapping). Initially, there were more forced trials than free trials to help them understand the task. Once the animals reliably differentiated between the low and high lottery choice trials, more free trials were added. The intermediate frequencies were added one by one, contingent upon consistent behavior in the free trials with existing frequencies. The lottery probability and/or the surebet magnitude were adapted to each animal so that their preferences could be reliably estimated. For example, if an animal chose the lottery too often, we would decrease the lottery probability. On average, mice took 129 ± 56 days to complete training and have their sessions meet the criteria of inclusion, whereas rats required only 90 ± 39 days.

Inclusion criteria for rodent sessions As described in the previous section, animals initially experienced two sounds mapped to one very bad ($EV = 0$) and one very good lottery. Animals that did not demonstrate a preference for the good lottery over the bad lottery were removed from the experiment. Additional lotteries were slowly added until animals chose between 5 or 6 lotteries each session. Animals whose behavior became unstable after the addition of new lotteries were reverted to fewer lotteries to re-stabilize and try again to increase the number of lotteries. In order for a session of animal behavior to be included into the dataset, it needed to have > 4 lotteries, > 40 choice trials, and at least 10% of choices for both lottery and surebet. In addition, we only included data from sequences of sessions where over half (in a 30 session window) would be included. In other words, if there was a single 'good' session in the middle of sequence of 20 sessions < 40 choice trials, that single 'good' session would not be included. We did not filter out sessions where choice was

uncorrelated with expected value. However, since progression through the training stages required choice to be correlated with expected value, each animal was effectively filtered by this feature. 169
170

Human risky choice task 171

In total, 119 human subjects (86 women, 29 men; age between 18 and 78) participated in an online study 172
hosted on Pavlovia. The subjects were recruited through NYU Psychology Department SONA paid subject 173
pool and this study was approved by the IRB of NYU Shanghai. The consent form indicated that each subject 174
will receive \$5 (US dollars) for their participation until the end of the session and each subject may receive a 175
bonus between \$5 and \$20. To facilitate training, subjects were given hints about the meaning of the sound in 176
the task: 177

“We would like you to play the game and earn coins. If you select the blue circle, coins you earn 178
are always the same. If you select the yellow circle, coins you earn will vary. After selection, press 179
white circle to collect coins ... You will hear a sound in the beginning of each trial. Pitch of this 180
sound varies with the number of coins. Rhythm of this sound varies with the chance to earn the 181
coins. Please make sure you are not muted. Press the spacebar to start the training trials. The 182
coins you earn will not be added to the total.” 183

Then subjects had 20-22 training trials and continued with the real task, where each coin was worth money 184
(the exact conversion rate differed from session to session since the distribution of the acquired coins for each 185
session was rescaled to resemble a \$5-\$20 range for the bonus payment, on average subjects in the nonverbal 186
task got bonus of \$12). There were three versions of the task during the initial session. Each session followed 187
the same timeline (FIGURE 2A). First, subjects heard the sound cue (2 seconds) and saw two circles (blue for 188
surebet and yellow for lottery option). Then, subjects made a choice and either got the constant reward for 189
sure if they chose the blue circle or a random draw determined whether they got the lottery reward if they 190
chose the yellow circle. At the same time, subjects saw a visual of money and money bag and heard rewarding 191
coin dropping sound, if they got reward in this trial. The choice set from the first version is in FIGURE 2C, 192
LEFT (22 training + 180 real trials). In the second version (compared to the first version), we added images of 193
coins of what a subject would have earned, even if he lost the lottery on a given trial. We also added six forced 194
trials in the beginning of the training phase. In the third version (compared to the first version), we changed 195
the choice set to the one in FIGURE 2C, RIGHT and added six forced risk trials in the beginning of the 196
training phase (20 training + 216 real trials). We followed up with 50 subjects who showed at least some 197
sensitivity in their choices, tested by a generalized linear model (GLM): 198

$$\text{choose_lottery} \sim \text{lott_mag} + \text{lott_prob} \quad (1)$$

where *choose_lottery* is 1 if lottery was chosen on a trial, *lott_mag* is the lottery magnitude and *lott_prob* is 199
the lottery probability on this trial. Sensitivity to stimuli was identified through both slope coefficients being 200
statistically significant at 0.05 level. The sensitivity to stimuli did not depend on the version of the task during 201
the first session. The two follow-up nonverbal sessions used the same design as in the third version (20 training 202
+ 168 real trials, a few subjects did not participate in the second follow-up session). None of the subjects 203
actions were considered as violations. Subjects could take as long as they needed to complete the session (i.e. 204
the same number of trials for each subject per session). A few subjects completed less trials than needed to 205
complete the session because of the lost internet connection. Most of the sessions took around or less than 30 206
minutes to complete. Only 36 of these subjects were sensitive to both reward and probability according to the 207
GLM comprised of trials from all sessions or at least the follow-up sessions. 208

Modeling Risky Choice 209

We considered three functional forms of risk, two noise specifications and three mixture formulations, resulting 210
in a total of 18 models. This section describes each model in detail. 211

History-independent models

212

As the lottery offer was independent on a trial-by-trial basis, we first developed models that only consider the current trial's information. An animal that understands the task perfectly should not rely on any past trials' information to guide their behavior.

213

214

215

The rho-sigma model The rho-sigma model models risk preference ρ as the exponent of the subjective utility curve:

216

217

$$U_L = V_L^\rho P_L \quad (2)$$

$$U_{SB} = V_{SB}^\rho \quad (3)$$

$$\hat{U}_L \sim \mathcal{N}(U_L, \sigma) \quad (4)$$

$$\hat{U}_{SB} \sim \mathcal{N}(U_{SB}, \sigma) \quad (5)$$

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (6)$$

(7)

where the expected utility of the lottery (U_L) and the utility of the surebet (U_{SB}) are computed from the lottery magnitude (V_L), lottery probability (P_L) and surebet value (V_{SB}). The internal noisy representations of the expected utility of the options (\hat{U}_L, \hat{U}_{SB}) are modeled as unbiased normal distributions with standard deviation σ . The probability of choosing lottery on this trial is the probability of \hat{U}_L larger than \hat{U}_{SB} , which can be solved analytically:

218

219

220

221

222

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (8)$$

$$= 1 - \Phi(0; V_L^\rho P_L - V_{SB}^\rho, \sqrt{2}\sigma) \quad (9)$$

where $\Phi(0; V_L^\rho P_L - V_{SB}^\rho, \sqrt{2}\sigma)$ is the cumulative Normal distribution with mean $V_L^\rho P_L - V_{SB}^\rho$, standard deviation $\sqrt{2}\sigma$ and evaluated at 0.

223

224

The kappa-sigma model The kappa-sigma model models risk preference κ as the degree of variance aversion:

225

226

$$\hat{U}_L \sim \mathcal{N}(P_L(V_L - \kappa\sqrt{\text{Var}(V_L)}), \sigma) \quad (10)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}, \sigma) \quad (11)$$

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (12)$$

where $\text{Var}(V_L)$ denotes the variance of the lottery option.

227

The rho-kappa-sigma model The rho-kappa-sigma model is the hybrid of the rho-sigma and kappa-sigma model. It models risk preference using two terms: ρ for the curvature of utility function, and κ for the degree of variance aversion:

228

229

230

$$\hat{U}_L \sim \mathcal{N}(P_L(V_L^\rho - \kappa\sqrt{\text{Var}(V_L^\rho)}), \sigma) \quad (13)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}^\rho, \sigma) \quad (14)$$

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (15)$$

The scalar models The scalar model only differs from its non-scalar counterpart in that the internal noise of expected utility scales with its magnitude, as in scalar utility theory (Kacelnik and Brito e Abreu, 1998).
The coefficient of variation γ in $\mathcal{N}(EU, \gamma EU)$ was estimated instead of σ in the non-scalar models.

Formally, for the rho-scalar model, the internal representation of expected utility is computed as:

$$\hat{U}_L \sim \mathcal{N}(P_L V_L^\rho, \gamma P_L V_L^\rho) \quad (16)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}^\rho, \gamma V_{SB}^\rho) \quad (17)$$

For the kappa-scalar model:

$$\hat{U}_L \sim \mathcal{N}(P_L (V_L - \kappa \sqrt{\text{Var}(V_L)}), \gamma P_L (V_L - \kappa \sqrt{\text{Var}(V_L)})) \quad (18)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}, \gamma V_{SB}) \quad (19)$$

For the rho-kappa-scalar model:

$$\hat{U}_L \sim \mathcal{N}(P_L (V_L^\rho - \kappa \sqrt{\text{Var}(V_L)}), \gamma P_L (V_L^\rho - \kappa \sqrt{\text{Var}(V_L)})) \quad (20)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}^\rho, \gamma V_{SB}^\rho) \quad (21)$$

The three-agent mixture models The three-agent model is a mixture model with three agents: the preference-expressing ‘rational’ agent, the habitual ‘lottery’ agent that always chooses the lottery, and the habitual ‘surebet’ agent. The final probability of choosing lottery is thus a weighted average of the three agents, each implementing a different behavioral strategy. Specifically, each agent outputs a probability of choosing lottery that makes up the probability vector \vec{P} , which is combined using their respective weights $\vec{\omega}$:

$$\vec{P} = [P_{rational}, 1, 0] \quad (22)$$

$$\vec{\omega} = [\omega_{rational}, \omega_{lottery}, \omega_{surebet}] \quad (23)$$

$$\sum \vec{\omega} = 1 \quad (24)$$

$$P(\text{Choose Lottery}) = \vec{P} \cdot \vec{\omega} \quad (25)$$

Below is a list of all the history-independent three-agent mixture models (mix). Note that they only differ in how the model specifies the rational agent.

- The mix-rho-sigma model
- The mix-kappa-sigma model
- The mix-rho-kappa-sigma model
- The mix-rho-scalar model
- The mix-kappa-scalar model
- The mix-rho-kappa-scalar model

History-dependent models

250

Although animals can be sensitive to a long reward history (cite) here, we only considered the effects of the $t - 1$ trial's outcome. It can be a 'lottery-win' – the lottery was chosen and it delivered, a 'lottery-lose' – the lottery was chosen and it did not deliver, and a 'surebet' – the surebet was chosen. To incorporate history effects into the three-agent mixture models, we allowed the mixture weights to vary depending on the previous trial's outcome. For example, $\omega_{lottery}^{SB}$ denotes the weight of the lottery agent after a previous surebet choice, and $\omega_{rational}^{lose}$ denotes the weight of the rational agent after a lottery-lose. Concretely,

$$\vec{P} = [P_{rational}, 1, 0] \quad (26)$$

$$\vec{\omega}_{win} = [\omega_{rational}^{win}, \omega_{lottery}^{win}, \omega_{surebet}^{win}] \quad (27)$$

$$\vec{\omega}_{lose} = [\omega_{rational}^{lose}, \omega_{lottery}^{lose}, \omega_{surebet}^{lose}] \quad (28)$$

$$\vec{\omega}_{SB} = [\omega_{rational}^{SB}, \omega_{lottery}^{SB}, \omega_{surebet}^{SB}] \quad (29)$$

$$\sum \vec{\omega}_{win} = \sum \vec{\omega}_{lose} = \sum \vec{\omega}_{SB} = 1 \quad (30)$$

$$P(\text{Choose Lottery}) = \begin{cases} \vec{P} \cdot \vec{\omega}_{win}, & \text{if lottery-win} \\ \vec{P} \cdot \vec{\omega}_{lose}, & \text{if lottery-lose} \\ \vec{P} \cdot \vec{\omega}_{SB}, & \text{if surebet} \end{cases} \quad (31)$$

Below is a list of all the history-dependent three-agent models:

251

- The history-mix-rho-sigma model 252
- The history-mix-kappa-sigma model 253
- The history-mix-rho-kappa-sigma model 254
- The history-mix-rho-scalar model 255
- The history-mix-kappa-scalar model 256
- The history-mix-rho-kappa-scalar model 257

Parameters summary

258

Parameter	Meaning	Lb	Ub
ρ	exponent on the expected utility function	0	None
κ	degree of variance aversion	None	None
σ	noise in the internal representation of expected utility	0	None
ω_{agent}	mixing weight of an agent in the mixture models	0	1
$\omega_{agent}^{outcome}$	mixing weight of an agent after a previous outcome in the history-mix models	0	1

Table 1. Summary of all parameters specified in our models.

Analysis

259

For all analyses, we excluded time out violation trials (where the subjects disengaged from the ports for more than 30 s during the trial) and trials with reaction time longer than 3 s. All analysis and statistics were computed in R (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria)

260

261

262

Efficacy analysis In this task, the rats were required to reveal their preference between the guaranteed and risky rewards. To understand how their behavioral performance improved over training, we evaluated the 'efficacy' of their choices following Constantinople et al.'s method. For each session, we computed the average expected value per trial of an agent that chose randomly, and a reward-maximizing agent that always chose the port with higher expected value. The average expected value per trial from each animal's choices (animal) relative to these lower (random) and upper (maximizer) bounds. It was computed as follows:

263
264
265
266
267
268

$$\text{efficacy} = 0.5 \frac{\text{animal} - \text{random}}{\text{maximizer} - \text{random}} + 0.5 \quad (32)$$

First-order stochastic dominance To compute average first-order stochastic dominance for each species, we counted the number of lottery choices that violated first-order stochastic dominance, i.e. choosing the lottery when the surebet magnitude is higher than or equal to the lottery magnitude. We then obtained the percentage by dividing the number of these trials by the total number of trials where the surebet magnitude is higher than or equal to the lottery magnitude.

269
270
271
272
273

Generalized linear (mixed-effects) models Generalized linear models (GLMs) and generalized linear mixed-effects models (GLMMs) were fit using the `lme4` R package (Bates et al., 2015). To produce psychometric curves as in FIGURE 1E-F and FIGURE 2D, we specified a GLM where the probability of a lottery choice was a logistic function of $EV_{\text{lottery}} - EV_{\text{surebet}}$. The expected value of lottery is the product of the lottery magnitude and lottery probability ($EV_{\text{lottery}} = P_{\text{lottery}} \cdot V_{\text{lottery}}$). Similarly, EV_{surebet} denotes the expected value of surebet, which is simply the value of surebet here ($EV_{\text{surebet}} = V_{\text{surebet}}$, since $P_{\text{surebet}} = 1$). In standard R formula syntax:

274
275
276
277
278
279
280

$$\text{choose_lottery} \sim \text{delta_EV} \quad (33)$$

where `choose_lottery` is 1 if lottery was chosen on a trial, `delta_EV` is $EV_{\text{lottery}} - EV_{\text{surebet}}$.

281

To test whether the outcome of the previous trial affected choice on the current trial, we first classified the previous trial's outcome into three categories: lottery-win, lottery-lose and surebet. If the previous trial was a violation, we considered that as a surebet choice. A GLMM was specified for each species:

282
283
284

$$\text{choose_lottery} \sim \text{delta_ev} + \text{prev_outcome} + (\text{delta_ev} + \text{prev_outcome} | \text{subjid}) \quad (34)$$

where `prev_outcome` is a categorical variable with three levels of previous outcome as above, and `subjid` denotes the subject ID of each animal.

285
286

To test whether the ρ and κ estimates were different across species, we specified a linear mixed model (LM) as follows:

287
288

$$\text{parameter} \sim \text{species} \quad (35)$$

where `parameter` is the MAP estimates of ρ or κ , `species` is a categorical variable with three levels: human, rat and mouse.

289
290

Modeling fitting Following modern statistical convention, we estimated the posterior distribution over model parameters with weakly informative priors using the `rstan` package (v2.21.2; Stan Development Team, 2020). `rstan` is the R interface of Stan (Stan Development Team, 2020), a probabilistic programming language that implements Hamiltonian Monte Carlo (HMC) algorithm for Bayesian inference. For models containing ρ , the prior over the utility exponent ρ was $\text{Lognormal}(\log(0.9), 0.4)$, a weakly informative prior that prefers ρ to be close to risk-neutral. For models containing κ , the prior was $\mathcal{N}(0, 2)$. The prior over noise σ was $\text{Gamma}(6, 3)$. The prior over the mixing weights $\vec{\omega}$ was a Dirichlet distribution with the concentration

291
292
293
294
295
296
297

parameter $\alpha = [6, 2, 2]$. The resulting $\omega_{rational}$ distribution was broad and had the mean of 0.6, both $\omega_{lottery}$
298
and $\omega_{surebet}$ distribution had the mean of 0.2. By attributing more weight to the rational agent over the
299
habitual agents, the prior reflected our selection of the experimental animals - only the ones with good
300
psychometric curves were included. Four Markov chains with 1000 samples each were obtained for each model
301
parameter after 1000 warm-up samples. The \hat{R} convergence diagnostic for each parameter was close to 1,
302
indicating the chains mixed well.
303

Synthetic datasets To test the validity of our model, we first created synthetic datasets with parameters
304
generated from the prior distributions described above. Each model was fit to its synthetic datasets and
305
overall, they were able to recover the generative parameters accurately (FIGURE S2). This assured that our
306
models had no systematic bias in estimating the parameters.
307

Model prediction confidence intervals To estimate the confidence intervals with model prediction as in
308
FIGURE 3B, we first generated a synthetic dataset with out-of-sample lottery magnitudes (incremented by 1).
309
After parameter sampling in each iteration (in the **generated quantities** block), the sampled parameters
310
were used to predict the choices given the synthetic offers. The resulting output is a **n_iter** x **n_lott_mag**
311
matrix, where **n_iter** is the number of iterations and **n_lott_mag** is the length of unique lottery magnitudes.
312
Finally, ± 80 , ± 95 and ± 99 confidence intervals for each offer were estimated by taking the respective quantiles
313
of the **n_iter** predicted choices.
314

Model comparison To assess model performance on unseen data, we performed 10-fold cross-validation of
315
the MCMC fits. For each fold, the model first fit on training set for that fold. We then computed (in the
316
generated quantities block) the log predictive densities by passing in the held-out data, using the posterior
317
draws conditional on the training data. As the training and testing data are independent, the log predictive
density coincides with the log likelihood of the test data. To evaluate the predictive performance, we computed
the expected log pointwise predictive density (ELPD) using the test data (Vehtari et al., 2017). As the
definition of ELPD incorporates the true generating process of prediction that is unknown, in practice, ELPD
is approximated by computing the log predictive density using draws from the posterior samples

$$\widehat{lpd} = \sum_{i=1}^n \log \left(\frac{1}{S} \sum_{s=1}^S p(y_i | \theta^s) \right), \quad (36)$$

where n is the number of test trials, θ^s is the s-th parameter sample from the posterior, and $p(y_i | \theta^s)$ is the log
315
predictive density of the i-th test trial computed using the s-th parameter sample. Intuitively, the closer ELPD
316
is to 0, the higher the model predictive accuracy.
317

Results

A cross-species paradigm of risky choice

Thirty-nine rats and sixty-four mice were trained on the risky choice task. On each trial, the animal chose
320
between a surebet and a lottery with fixed probability and auditory-guided magnitude (FIGURE 1A-B, see
321
task details in Methods). The data in this manuscript includes 14,591 behavioral sessions from rodents (8,359
322
from mice, 6,232 from rats), with a total of 1,509,917 choice trials (821,657 from mice, 688,260 from rats). On
323
average, mice performed 91.5 ± 33.3 (mean and standard deviation) choice trials per session, and a rat
324
performed 96.9 ± 35.0 choice trials per session (FIGURE 1C). A trial was considered a violation if the animal
325
failed to poke into central 300 s after trial start, or it did not make a choice 30 s after fixation, thereby ‘opting
326
out’ this trial. The percentage of violation trials is a proxy of the animal’s productivity: the higher the
327
violation rate, the fewer trials they completed. On average, mice opted-out $5\% \pm 4$ and rat opted-out $4.9\% \pm 2$
328
of all trials in a session, suggesting that both species were quite motivated to perform the task (FIGURE 1D).
329

The animals learned the task after extensive training. Rats required fewer training days than mice to
330
approach the efficacy of an expected value maximizer than mice (mice: 129 ± 56 days, rats: 90 ± 39 days,
331

FIGURE S1). The animals' choices were largely consistent with a utility-maximizing strategy: the average frequency of first-order stochastic dominance violations (i.e. choosing the lottery when surebet magnitude was equal or higher) was 23.6% in mice and 14.9% in rats. While these proportions may seem high for a typical verbal study of risk-attitudes, they are similar (or lower) than we saw in our human subjects (see next paragraph). Although most animals increased the proportion of lottery choices monotonically with increasing expected value (FIGURE 1E-F), some animals did not (e.g. FIGURE S3 1307). Note, non-monotonic utility functions are a feature of variance-aversion in 'willingness-to-pay' models (Tobler and Weber, 2013). The animals further demonstrated learning of the task by opting out more from the trials with smaller lottery magnitudes (FIGURE 1G-H). There was no benefit to be gained from opting-out – completing the trial would result in some reward and faster progression to the next trial. This is consistent with results from Constantinople et al. (2019b), where rats chose between a safe and risky option with varying magnitudes and probabilities.

The human task closely mirrored the rodent task, where subjects made a choice between a surebet and a lottery option, whose magnitude and probability was signaled by a sound cue (FIGURE 2). We included 88 sessions (14,180 trials) from 36 subjects; all of them were sensitive to both lottery magnitude and probability (tested by a GLM). The task was structured such that a risk-neutral subject would choose the lottery 50% of the time, and we found that on average, our human subjects chose lottery 52% of the time. This is somewhat surprising, as human subjects are typically found to be risk-averse (23and Me Research Team et al., 2019; Tymula et al., 2013). We included lotteries that were strictly dominated by the surebet – in other words, no rational subject (in the Von Neumann and Morgenstern, 1953, sense of rational) would choose the lottery on those trials. The frequency of first-order stochastic dominance violations was 28% on average; this is higher than what was found in tasks where the offers were described in text: Tymula et al. (2013) reported ~5% for the same age group. Our subjects significantly decreased the number of dominance violations towards the third session from 34% to 22% (permutation test, $p < 0.001$), reaching levels smaller than those observed for older adults (Tymula et al., 2013). These levels of violations are consistent with previous investigations of nonverbal economic decision-making (Lukinova et al., 2019).

Systematic modeling of risky choice

We considered three functional forms of risk. Firstly, we can model risk-tolerance ρ as the curvature of a power-law utility function, $U = V^\rho$ (FIGURE 3A, TOP), with $\rho < 1$ indicating risk-aversion and $\rho > 1$ indicating risk-seeking. Secondly, risk-tolerance can be formulated as the degree of aversion to the variance of the expected distribution of returns (κ , FIGURE 3A, BOTTOM), as in the mean-variance models. In this case, $\kappa > 0$ indicates variance-aversion and $\kappa < 0$ indicates variance-seeking. Last but not least, we can combine the first and second form to capture both the curvature of utility function and the degree of variance aversion, as these two forms are not mutually exclusive. We then considered two specifications of noise. First, a single parameter σ can represent noise in the utility function as in $N \sim (EU, \sigma)$, where EU is the expected utility of lottery or surebet (FIGURE 3B, TOP). This form of noise assumes that animal has a noisy representation of its utility function, and the noise is constant irrespective of the perceptual input. Second, we can also model noise according to scalar utility theory (Kacelnik and Brito e Abreu, 1998), in which the noise follows Weber's law – it scales with utility magnitude $N \sim (EU, \gamma EU)$ and γ represents the coefficient of variation (FIGURE 3B, BOTTOM).

Finally, we considered three formulations of heuristic strategy mixtures. The first models subjects as a single utility-maximizing agent. Referred as the 'rational' agent, this agent can express any combination of the risk forms and noise specifications. This is the standard approach in most human and non-human primate studies (e.g. Farashahi et al., 2018; Stauffer et al., 2014; McCoy and Platt, 2005). The second formulation assumes that the final choice is a weighted outcome of three agents: a utility-maximizing agent, a habitual 'lottery' agent that always chooses lottery and a habitual 'surebet' agent. The relative influence of the agents is controlled by their mixing weights ω , where $\sum \vec{\omega} = 1$. The choice on each trial is thus a weighted outcome of the 'votes' of three agents, each implementing a different strategy (FIGURE 3C, LEFT). It has been well documented that rodents have a tendency to repeat the choice that led to rewards on the previous trial and switch away from the choice that led to no reward, known as 'win-stay lose-shift' (Stopper and Floresco, 2011; Marshall and Kirkpatrick, 2013; Constantinople et al., 2019b). As such, the last kind of mixture model

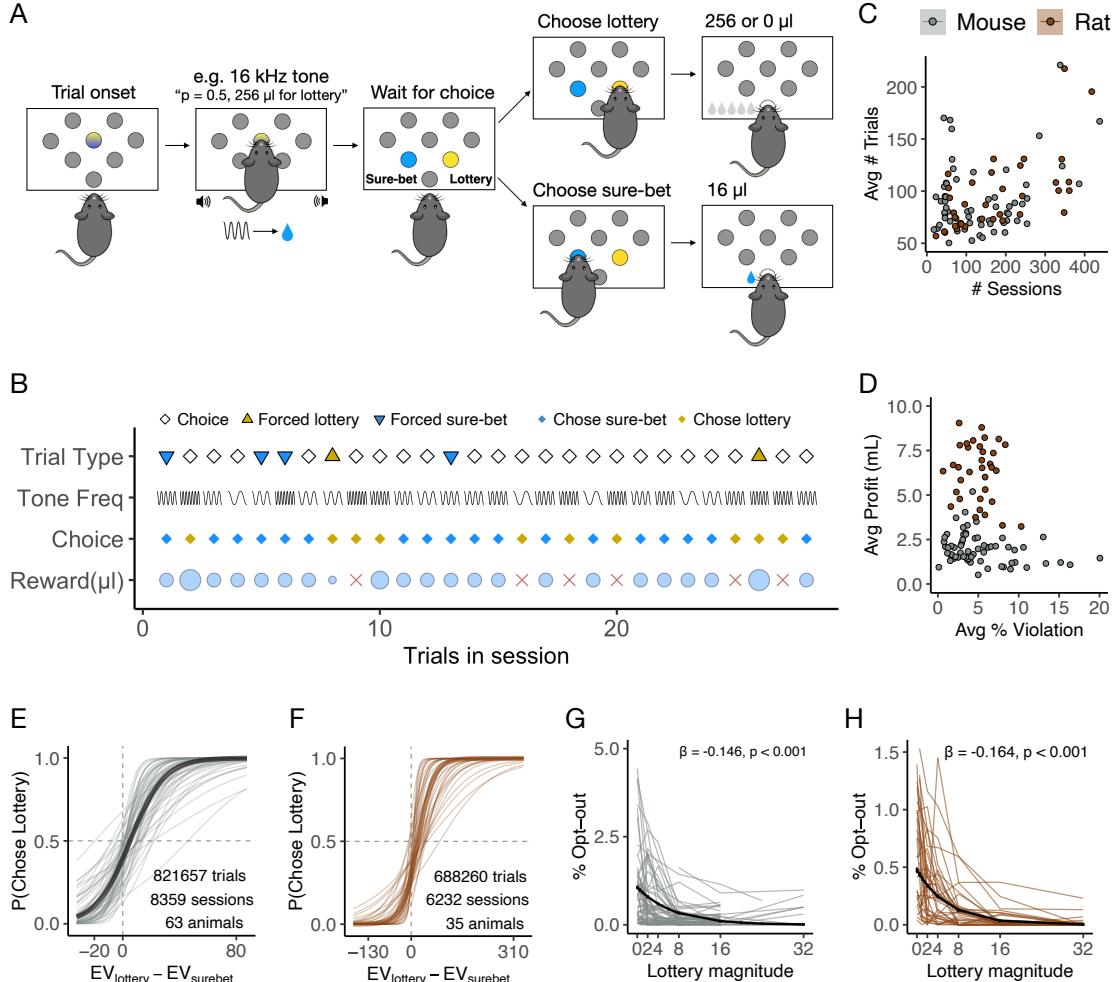


Figure 1. The risky choice task and animal behavior. **A.** Schematic of the risky choice task. Each trial began with the onset of central LED, which cued the animal to poke into the center port and hold there for 1s. An auditory cue was played, indicating the magnitude of the lottery. After 1s, the animal withdrew from the center port and made a choice poke into the surebet or the lottery port. A surebet choice delivered a small and guaranteed reward, whereas a lottery poke gave either 0 or the cued magnitude. The probability of a ‘win’ was fixed within a session. **B.** An example trial sequence. For trial type, white diamond, yellow triangle and blue triangle represents the choice trial, forced lottery trial and forced surebet trial, respectively. The sine-waves in the ‘tone freq’ row symbolize the frequency of tones played. Choices are marked in diamonds, with yellow for lottery and blue for surebet. The reward received (μl) on each trial is shown in light blue circles, whose size represents the relative amount. The red cross represents a lottery lose with no rewards. **C.** The number of sessions and the average number of trials per session for mice ($n = 64$, in gray) and rats ($n = 39$, in brown). Each dot represents one subject. **D.** The average percentage violation (opt-out) per session and average water earned (mL) per session. The base reward volume was higher for rats than mice due to their larger body mass. **E.** Task performance from all mice. The probability of choosing lottery is plotted as a function of the expected value of lottery minus the expected value of surebet ($V_{\text{lottery}} P_{\text{lottery}} - V_{\text{surebet}}$), where V represents μL of water. The lines were generated by a generalized linear model, the thin gray lines are fit to data from each mouse, the thick gray line fit to data from the mouse population. **F.** Same as **E** from all rats. **G.** The opt-out behavior in mice. The lines were generated by a generalized linear model, the thin lines fit to data from each animal, the thick line fit to data from the population. β denotes the main effect of lottery magnitude on % opt-out. **H.** The opt-out behavior in rats.

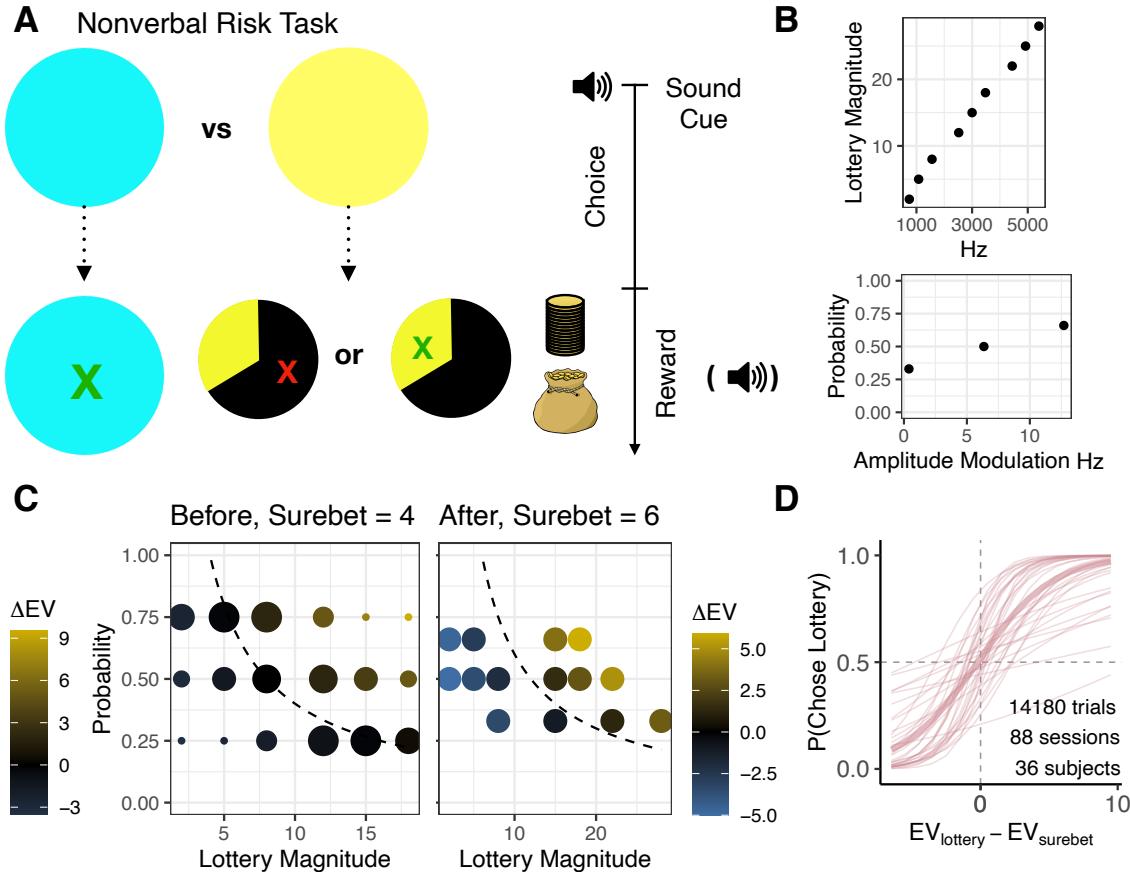


Figure 2. The human risk nonverbal task and behavior. **A.** Schematic of the choice and reward screens. Left: the blue circle represents the surebet (constant), the yellow circle represents the lottery, whose magnitude (lottery magnitude and lottery probability were indicated via a sound cue). Timeline of a single trial of the risk nonverbal task: subjects heard the sound cue (2 seconds) and saw two circles, subjects made a choice and either got the constant reward for sure if they chose the blue circle or a random draw determined whether they got the lottery reward if they chose the yellow circle. **B.** Sound map from the final version of the task. Lottery magnitude was indicated by tone frequency (Hz), lottery probability was indicated by amplitude modulation (Hz). **C.** Left: choice set from the first version. The size of the circle visualizes the number of trials in each conditions (18 unique lotteries = 6 magnitudes x 3 probabilities, each repeated from $n = 4$ to $n = 16$). Right: choice set from the final version. Each offer (14 unique lotteries: lottery probabilities: 0.33, 0.5, 0.66 and magnitudes: 2, 5, 8, 12, 15, 18, 22, 25, 28) was repeated exactly $n = 8$ times. The dashed curve is the indifference line ($\Delta EV = 0$), indicating no difference between expected values of the lottery and the surebet. The lottery options were designed to be roughly symmetric around the dashed curve, such that a risk-neutral subject would choose the lottery 50% of time. **D.** Task performance from the human population. Same as FIGURE 1E & F. Only now V represents number of coins.

considers the effects of trial history, specifically, whether the previous trial's outcome affects the weights of the three agents. This can be formulated by allowing $\vec{\omega}$ vary depending on the outcome of the previous trial, be it a lottery-win, lottery-lose or surebet (FIGURE 3C, RIGHT). 383
384
385

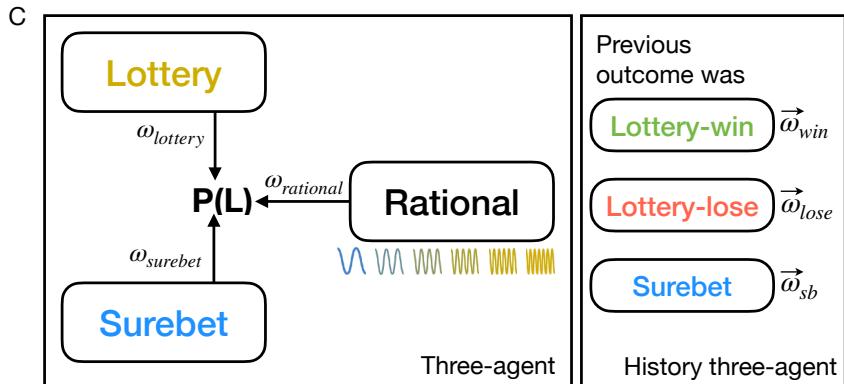
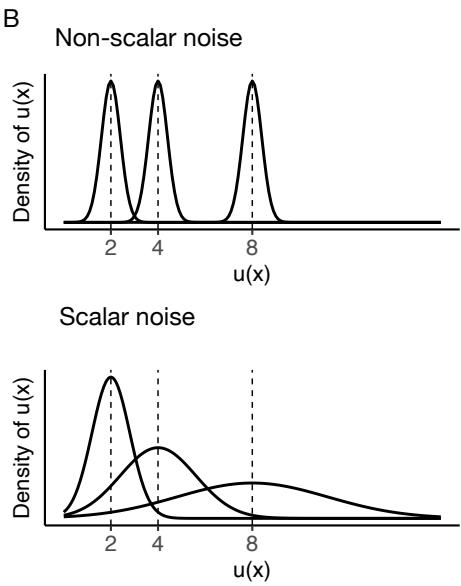
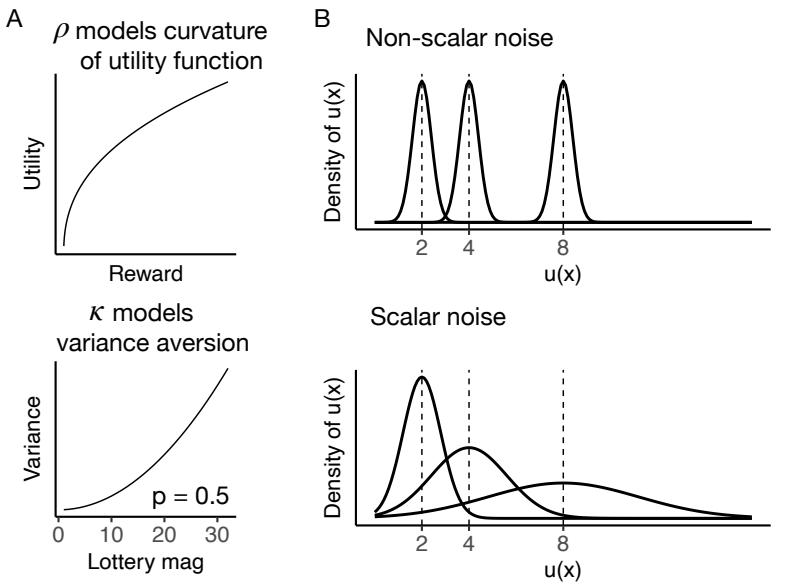
With three risk forms, two noise specifications and three mixture formulations, we constructed 18 models of risky choice in total (FIGURE 3D, see all models in Methods). For naming conventions, 'history-mix-rho-kappa-sigma' refers to a model that uses both ρ and κ as functional forms of risk-tolerance, a standard σ noise specification and the history mixture formulation. Take another example, 'rho-scalar' refers to a model using ρ as risk form, a scalar noise specification and no mixture. All models have been validated to be able to recover generative parameters from synthetic data (FIGURE S2). We estimated the joint posterior over the parameters for each subject separately using Hamiltonian Monte Carlo sampling in Stan (Carpenter et al., 2016). 386
387
388
389
390
391
392
393

Model-based estimation of risk attitudes 394

To quantify risk attitude for each species on the population level, and to facilitate comparison with previous findings, we first describe the fits using the 'rho-sigma' model, a standard model used in economic research. The rho-sigma model produced adequate fits to the data (see examples in FIGURE 4B, more examples in FIGURE S3, S4, S5). All mice were characterized to be risk-averse ($0.62[0.18, 1.05]$, median and 95 % C.I. of concatenated posteriors), exhibiting concave utility functions (FIGURE 4A LEFT). Similarly, all rats were also estimated to be risk-averse by this model ($0.61[0.37, 0.96]$, FIGURE 4A MIDDLE). The human population, on the other hand, was comprised of both risk-averse and risk-seeking individuals, with the population median being close to risk-neutral ($0.91[0.32, 1.44]$, FIGURE 4A RIGHT). As can be observed from the distribution of posterior ρ samples (FIGURE 4C, D), humans are more risk-seeking than rodents on the population level ($\beta_{mouse} = -0.29 \pm 0.04, p < 0.001, \beta_{rat} = -0.29 \pm 0.04, p < 0.001$; Linear model of MAP ρ estimates, see Methods for details). 395
396
397
398
399
400
401
402
403
404
405

The concave utility functions found in rats are consistent with previous results (Constantinople et al., 2019b). To our knowledge, this is the first dataset describing the curvature of utility functions in mice. We found that mice and rats did not differ significantly (when parameterized by the rho-sigma model). Humans exhibited both concave and convex utility functions, different from previous literature which generally finds mostly concave functions (Holt et al., 2002; Tymula et al., 2013). Nonetheless, the risk attitude estimated from a non-verbal task are expected to have higher variability than those from a verbal task (Ludvig and Spetch, 2011; Wu et al., 2011). 406
407
408
409
410
411
412

Through visual inspection of the psychometric curves of animals and the rho-sigma fits of the data, it is clear that this simple model missed important aspects of the data (e.g. Rat 2116 in FIGURE 4B). As a comparison, we examined the estimates of ρ and the quality of fits from the model that considers all the elements – the 'history-mix-rho-kappa-sigma' model, and found that the estimates of ρ shifted substantially (FIGURE 5A). The median ρ increased for all three species: mice ($0.84[0.13, 1.86]$), rats ($0.80[0.15, 1.69]$), and humans ($1.00[0.08, 4.23]$). The increase in ρ was due to the addition of variance-aversion, κ , which, like ρ , influences the indifference between lotteries and the surebet (κ median & 95%CI: mice $0.20[-0.72, 0.97]$, rats $0.19[-0.72, 0.92]$, humans $0.01[-1.17, 0.78]$; FIGURE 5C, D). In addition, the heuristic mixture formulation enabled the model to capture stimulus-independent biases, fitting subjects like 2116 better while also more accurately estimating its utility function (FIGURE 5B). Moreover, the incorporation of trial-history-dependent parameters allowed the history effects in subjects like 1368, 1278 and 2181 to be estimated (FIGURE 5B). On the population level, the MAP ρ estimates are not different between mice and humans ($\beta_{mouse} = -0.14 \pm 0.11, p = 0.213$) and marginally different between rats and humans ($\beta_{rat} = -0.27 \pm 0.13, p < 0.05$). The MAP estimates of κ are highly different between mice and humans ($\beta_{mouse} = 0.30 \pm 0.10, p < 0.01$), but not so between rats and humans ($\beta_{rat} = 0.22 \pm 0.11, p = 0.054$; FIGURE 5C,D). Overall, these results suggest that once we isolated the history-dependence and stimulus-independent biases, the distribution of ρ and κ across species become more overlapped. 413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429



D

Risk forms	Noise	Mixture	
- rho	- non-scalar	- no mixture	$3 \times 2 \times 3$
- kappa	- scalar	- three-agent	18 models
- rho-kappa		- history three-agent	

Figure 3. Systematic modeling of risky choice. **A.** Functional forms of risk. Top: ρ models the curvature of utility function as in standard economic models. Bottom: κ models the subject's aversion to variance as a function of lottery magnitude and probability (here $p = 0.5$ only). **B.** Noise specification. Top: the subject has a constant noise in utility representation. Bottom: the subject has a scalar noise, which scales the magnitude of expected utility. **C.** Mixture formulation. Left: the three-agent mixture. The probability from the rational agent (can be of any risk forms and noise specification) is mixed with the output from a heuristic lottery and surebet agent with their respective weights ω , where $\vec{\omega} = 1$. Right: the history three-agent mixture. To include trial history effects, we allowed the weights vector $\vec{\omega}$ to vary depending on the previous trial's outcome. **D.** Combining three functional forms of risk, two noise specification and three mixture formulation, we have $3 \times 2 \times 3 = 18$ models of risky choice.

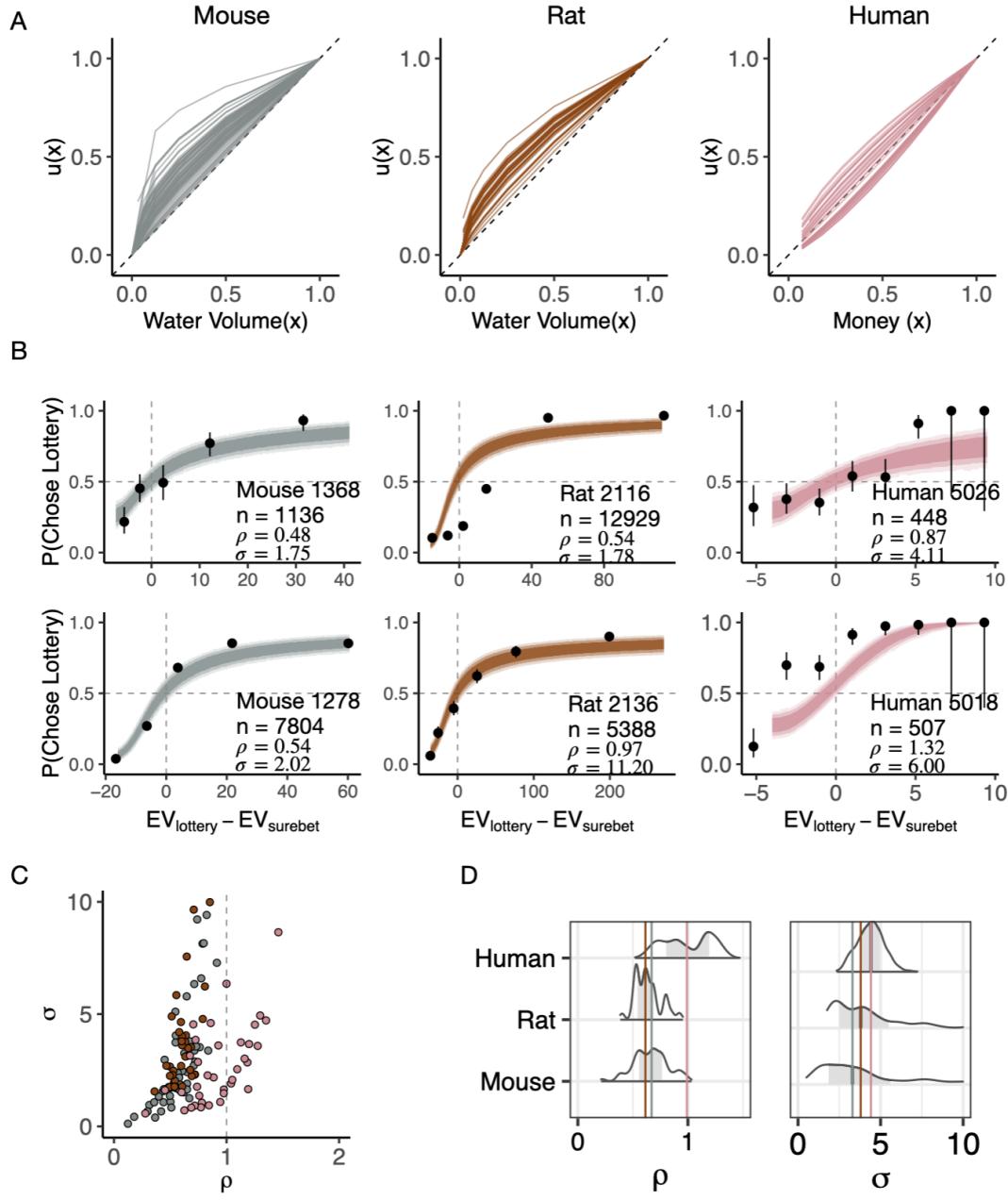


Figure 4. Fits by the rho-sigma model. **A.** The subjective utility functions for each subject computed using the maximum a posteriori (MAP) ρ estimation, normalized by the maximum reward amount. **B.** Example subjects from the mouse (gray), rat (brown) and human (pink) population. The circles with error bars are the mean $\pm 95\%$ binomial confidence intervals of binned choices. The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively. **C.** Distribution of ρ and σ estimates. Each point is a MAP estimate for one subject, colored by species. **D.** Density plots of combined posterior samples ($n = 100$ for each subject). The light gray shaded area marks the 80% interval of the posterior estimate. The outline of the distribution extends to the 99.99% interval. The colored lines are the posterior distribution medians to ease comparison across populations.

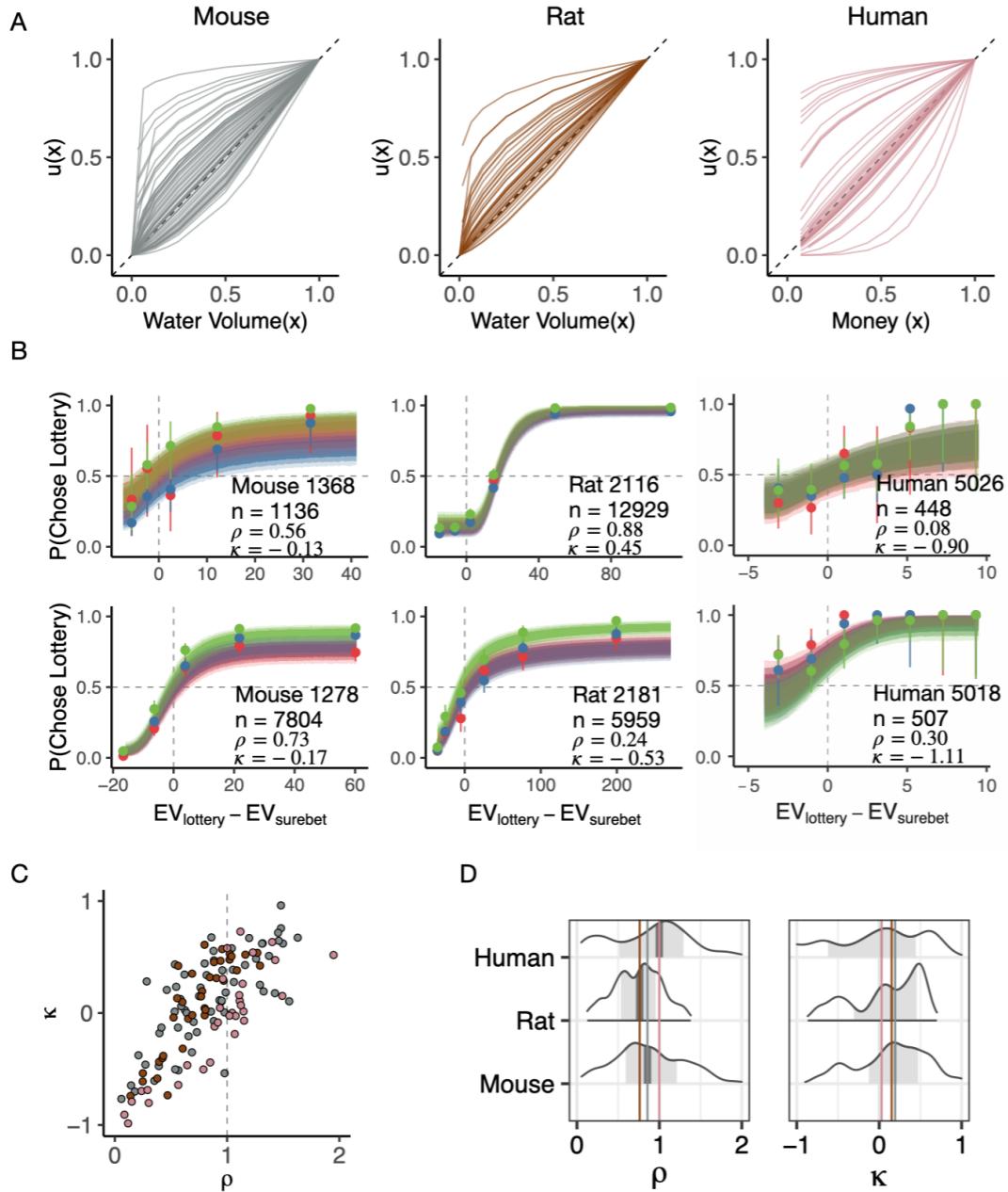


Figure 5. Fits by the history-mix-rho-kappa-sigma model. **A.** Same as FIGURE 4A. **B.** Same as FIGURE 4B except that the trials were plotted based on the previous outcome: lottery-win in green, lottery-lose in red, surebet in blue. **C.** Distribution of ρ and κ MAP estimates. **D.** Same as FIGURE 4D.

Animals are less utility-maximizing than humans

In the section above, we presented results from one of the smallest models (rho-sigma) and one of the largest models (history-mix-rho-kappa-sigma) to demonstrate that the choice of function form guides the inferences one might make about similarities and differences across species. To systematically compare models for each individual and the population overall, we performed 10-fold cross-validation using the 18 models on each subject's dataset. The expected log posterior density (ELPD) across all held-out data was used as the metric for model comparison, with ELPD closest to 0 indicating the best fit (Vehtari et al., 2017, see details in Methods). Although the best-fitting model varied considerably from individual to individual (FIGURE S12, S13, S14), differences emerged between humans and rodents on the population level. For both mice and rats, the most frequent best-fitting model was 'history-mix-rho-kappa-sigma' (best model in 28.6% mice, 37.1% rats), suggesting that rodents relied on heuristics as part of their strategy and were sensitive to the previous trial's outcomes, consistent with previous reports (FIGURE 6A, B). Their noise was best captured by a simple σ specification, suggesting that the internal noise is invariant to lottery magnitudes. Furthermore, the top-5 best-fitting models, amounting to 82.3% total for mice and 88.3% for rats, were all 'history-mix' models (FIGURE 6B). We specified a GLMM and found significant main effects of previous trial's outcome in mice ($\beta_{lottery-win} = 0.42 \pm 0.04, p < 0.001$; $\beta_{lottery-lose} = 0.24 \pm 0.04, p < 0.001$) and rats ($\beta_{lottery-win} = 0.44 \pm 0.04, p < 0.001$; $\beta_{lottery-lose} = 0.15 \pm 0.03, p < 0.001$), confirming the model comparison results. However, the ELPD error bars considerably overlapped among the top models for many subjects (e.g. FIGURE S12 1108, 1121, 1297, S13 2105, 2119, 2136), and several models performance was virtually indistinguishable by eye (FIGURE 6C). As such, we conclude that a class of preferred models was found for rodents and it contains heuristic strategy mixture and trial-history parameters.

In contrast, the most frequent best-fitting model for humans was 'rho-kappa-sigma' (19.4%, FIGURE 6A, B), indicating that these subjects were utility-maximizing without any stimulus-independent bias and history effects. The second most frequent best-fitting model was 'mix-rho-kappa-sigma' (16.3%), indicating that a heuristic strategy was adopted by these participants. In contrast to rodent model comparison results, all the 'history-mix' models produced worse fits than other models, and none of them (except in subject 5100) was elected to be the best-fitting model for any human subject. This is consistent with GLMM results that no significant history effects were observed in humans on average ($\beta_{lottery-win} = 0.01 \pm 0.09, p = 0.892$; $\beta_{lottery-lose} = -0.09 \pm 0.09, p = 0.333$). Nonetheless, as the ELPD errorbars from different models heavily overlapped in most human subjects (FIGURE S14), performance of the reported 'best-fitting' models were in fact indistinguishable from many other models FIGURE 6C). We thus emphasize that no definitive best model was found for humans on the population level. Overall, these results suggest that the key differences between rodents and humans are the use of heuristic strategies and dependence of trial history.

Estimation of ρ is model-dependent

Our model comparison analysis found that some classes of models were better than others while failing to definitively identify a best model for each subject. Can we still learn something from the estimation of risk attitudes from different models? To compare the estimation of ρ from all the models containing ρ ($n = 12$), we first extracted the MAP ρ estimate from each model for each individual, and then performed a rank correlation test on the model estimates for each pair of ρ -containing models. The resulting correlation matrices revealed interesting patterns within and between species (FIGURE 7). The most conspicuous finding was that ρ estimates from models fitting ρ and κ simultaneously had low to negative correlations with ρ estimates from models only containing ρ (FIGURE 7, BLUE SQUARE). This was true for all three species, albeit it was more pronounced in rats and mice than humans. The effect is unsurprising: in models with either ρ or κ (but not both), that parameter strongly influences the indifference point of the psychometric curve. When both parameters are included, they trade off (FIGURE 5C).

The next finding was that while a high correlation of ρ was observed within models only estimating ρ in humans (all $r > 0.84$), it was not the case with mice (smallest $r = 0.09$, mix-rho-scalar vs. rho-sigma) or rats (smallest $r = 0.52$, history-mix-rho-scalar vs. rho-sigma). Two things were inferred from this. First, results from humans suggest that our Bayesian models were well-specified – in the absence of heuristic mixture and history effects, the priors helped the models behave well even with extraneous parameters (FIGURE 7, HUMAN,

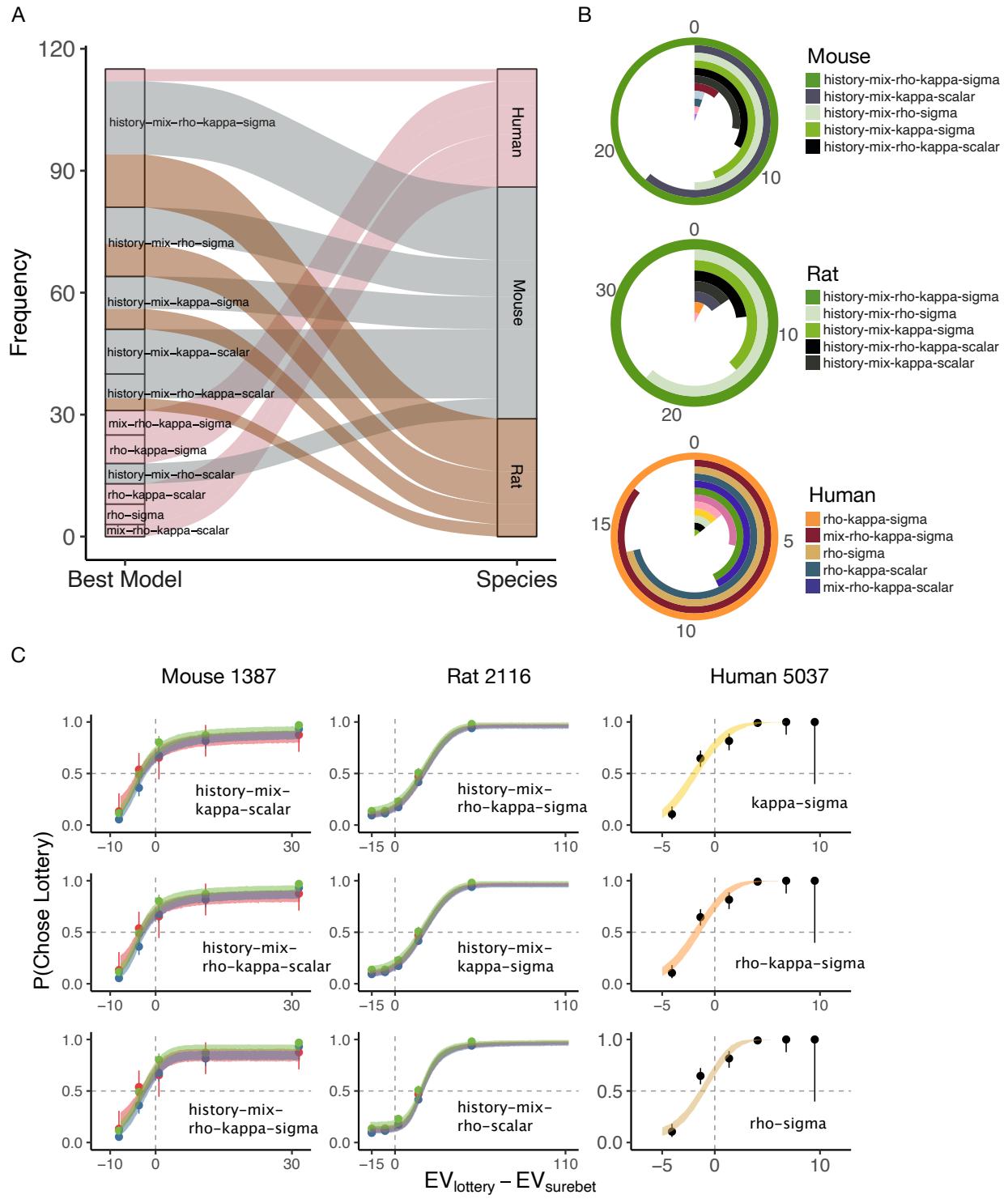


Figure 6. Model comparison with 10-fold cross-validation. **A.** Alluvial plot showing the frequency of the best-fitting model for each species. The width of the stratum represents the number of subjects belonging to this category. **B.** Circular bar plot showing the proportion (%) of the best-fitting model for each species, colored by model identity. The 5 most frequent best-fitting models are shown in legend. For the full legend, see FIGURE S5. **C.** Predictions from the top 3 best-fitting models of three example subjects. If the model includes history terms, the psychometric curves are plotted separately by the previous trial's outcome: green = lottery-win, red = lottery-lose, blue = surebet. The ribbons were generated by the fitted parameters with $\pm 80\%$ confidence interval.

YELLOW SQUARE VS. RED SQUARE). Second, for rodents (especially mice), the model comparison demonstrated that both required heuristic mixture and history parameters to describe the data. Since the correlation of ρ across models with and without these elements is low (ρ (FIGURE 7, MOUSE AND RAT, YELLOW SQUARE VS. RED SQUARE), we infer that the inclusion of these factors is important for correctly estimating ρ . It reaffirms our earlier conclusion that the heuristic strategy and history effects must be considered when modeling rodent behavior. Lastly, it appears that noise specification had little impact on ρ estimation when other model properties were matched. For example, ρ estimates from rho-scalar had a high correlation with that from rho-sigma in humans ($r = 0.95$), mice ($r = 0.89$) and rats ($r = 0.89$). Interestingly, ρ estimates from history-mix-rho-scalar were highly correlated with that from history-mix-rho-sigma in human ($r = 0.96$) and rats ($r = 0.97$), but less so in mice ($r = 0.72$). Overall, the correlation matrices help visualize how the estimation of ρ depends on the model as well as behavioral tendencies in the dataset.

480
481
482
483
484
485
486
487
488
489
490

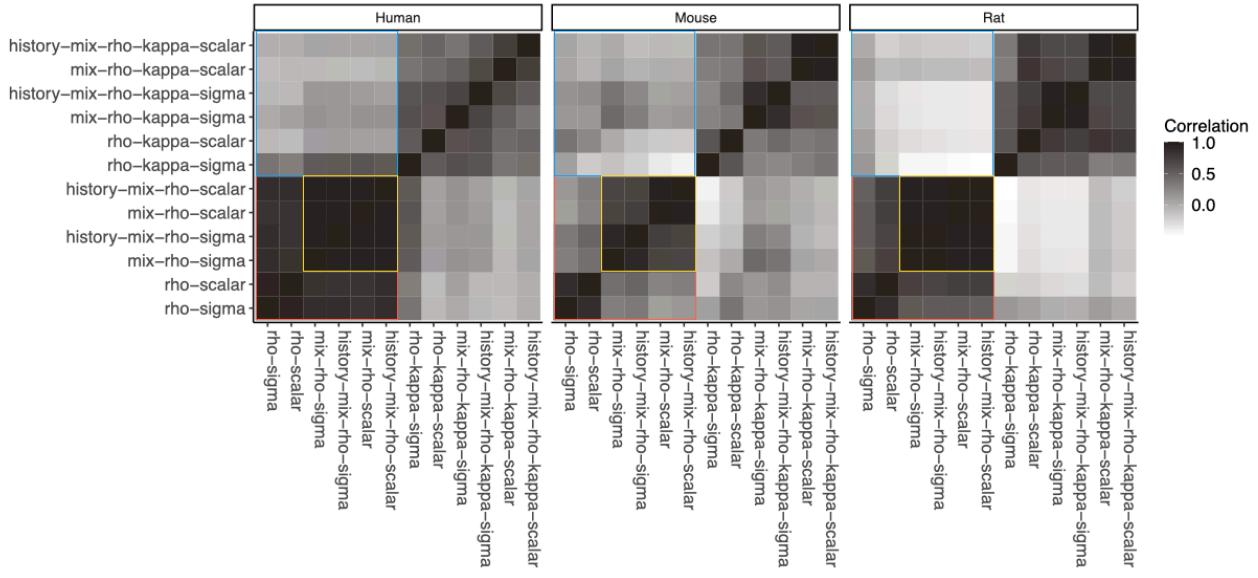


Figure 7. Species correlation of ρ estimates from all the models containing ρ . MAP ρ estimate from each model was obtained for each individual. Spearman's rank correlation test was performed between each pairwise model's estimation.

Discussion

491

Risky choice has been widely researched in humans and non-human animals. However, differences in task design render direct comparison of the results difficult. We developed a cross-species paradigm, in which the subjects chose between a surebet and a lottery option whose magnitude (and probability) was signaled by auditory cues. Mouse, rat and human subjects learned the sound mapping and most performed the task in line with utility-maximization. To decompose distinct elements of risk-tolerance, we developed a total of 18 models that considered heuristic mixture, history effect, noise specification as well as different functional forms of risk as in expected utility theory and mean-variance models. While human subjects were utility-maximizing, rodent behavior was better fit by a class of models that include heuristic mixture and history dependence. Using the best-fitting model overall, distributions of ρ and κ became more overlapping across species, and each population was comprised of risk-averse and risk-seeking individuals. Moreover, we showed that the estimation of ρ depended on model specification, suggesting that the choice of model is essential for the accurate construction of utility function. These results provide support for using mice and rats to examine the neurobiology of risk-tolerance, while caution against the use of overly simplified models which ignore certain behavioral tendencies in rodents.

Our results are consistent with that from Constantinople et al. (2019b), where rats were found to exhibit concave utility functions on average (FIGURE 5A). The median curvature of utility function was estimated to be 0.54, slightly smaller than what was estimated in our rats ($\rho = 0.80$). Importantly, their reference-point model included lapse and trial-history parameters, thereby making the ρ estimates comparable. In their study, rats showed win-stay-lose-shift behavior consistent with previous literature. Curiously, our rats were behaving in a way that can be described as ‘win-stay-lose-stay’, that they chose the lottery more after either a lottery-win or lottery-lose compared to post-surebet trials (GLMM results). It could be due to their tendency to persevere its previous choice, in which case would result in a pattern resembling win-stay-lose-stay. The influence of past information on current choice, even when it interferes with the optimal policy, has been observed in humans (Urai et al., 2017; Allefeld et al., 2013; Abrahamyan et al., 2016), monkeys (Gold et al., 2008; Padoa-Schioppa, 2013), and rodents (Hwang et al., 2017; Sul et al., 2011; Akrami et al., 2018; Morcos and Harvey, 2016). Such prevalent history-dependency suggests that tracking the past choice-outcome relationships to form a subjective bias is a fundamental aspect of decision-making. As animals evolved to exploit long-term regularities occurring in nature, the persistence of such behavioral disposition is unsurprising in an artificially random environment. We also demonstrated that the animals were variance-averse (median $\kappa > 0$ for both mice and rats). As far as we are aware, this is the first qualification of variance-aversion in mice and rats on a population scale. Nonetheless, we only included one lottery probability for each animal unlike in Constantinople et al. (2019b). The inclusion of more distinct lottery probabilities will yield a better estimate of κ , but it may prolong the task training time.

We found the human population exhibited both concave and convex utility functions from the nonverbal task, contrary to the general consensus that humans have concave utility functions (Holt et al., 2002; Haushofer and Fehr, 2014). Increased risk-seeking might be due to the ‘peanuts effect’, i.e. subjects choose lotteries more often in gambles with small stakes (Markowitz, 1952; Weber and Chapman, 2005). In each trial of the human nonverbal experiment, subjects played for coins, where each coin on average was about 0.004 US Dollars ($\sim \$12$ (bonus) / 200 (trials) / 15 coins (lottery reward)). With maximum number of coins being 28 = \$0.112 per trial, these were indeed small stakes. However, if subjects were playing for ‘peanuts’ then they would all show risk-seeking behavior or at least be more risk-seeking in the follow-up sessions, which was not the case. Instead, we argue that the mixed risk attitude is due to nonverbal nature of our task. We saw the increased range of risk preferences (extended more towards risk-seeking extreme) in the nonverbal task (stable across sessions) compared to what we expect from a standard verbal risk task (Holt et al., 2002; Tymula et al., 2013). This is in line with findings of Wu et al. (2011), who found the range of power utility parameters to increase in the motor domain compared to the classical verbal risk task. Although there are just a few studies that utilize nonverbal tasks for studying decisions under risk (Wu et al., 2011; Balci et al., 2009; Hanks and Summerfield, 2017; de Gee et al., 2018), studies that address the description-experience gap in risky choice abound (Hertwig and Erev, 2009). Increased risk-seeking was found for experienced gambles of equally probable outcomes relative to described ones, and importantly, persisted after extensive training (Ludvig and Spetch, 2011; Heilbronner and Hayden, 2016). Another explanation for risk-seeking from the literature is the repeated

gambles effect, that preferences in repeated gambles often move towards more risk-seeking (Samuelson, 1963; Lopes, 1981; Klos et al., 2005). Although we cannot rule out such effect due to our experiment design – our subjects made choices on whether to accept the lottery or to choose the surebet based on randomly ordered eight repetitions of fourteen unique lotteries – rarely in real life we are dealing with a single decision under risk.

The nonverbal task was specifically designed to bridge the gap between animal and human studies – both humans and rodents were choosing from experienced gambles. Nevertheless, two major differences exist. First, we used money as reward for humans as opposed to water for water-deprived animals. Although evidence suggest various reward prospects are represented on a common scale (Levy and Glimcher, 2012; Chib et al., 2009; Hare et al., 2008), differences were found in the representation of abstract secondary rewards compared to primary rewards (Sescousse et al., 2013). The extent to which the nature of rewards interacts with risk attitude remains unclear and represents an interesting research topic. Second, although the meaning of the sound was instructed to the subjects prior to their first session, they only had ~ 20 training trials to learn the mapping. This contrasts with our animals who had to learn the map without any explicit instruction, but their extensive training should make up for it. Thus, the nonverbal task is a mixture of making choices from instructed learning and experience. The finding that some humans (e.g. FIGURE S14 5064, 5059, 5100) were best fit by a mixture model suggests the use of heuristics. Robust heuristics rather than optimal inferences are often adopted under uncertainty, which was probably what the task appeared to these subjects before they fully understood the sound cues (Neth and Gigerenzer). As such, some subjects may be choosing without complete information of the options, thereby exhibiting reliance on heuristic strategies (Pisupati et al., 2021).

Lastly, we wish to emphasize that we only applied 18 risky choice models, whose total number is infinite in theory. One leftout class is the reference-point models from prospect theory (Kahneman and Tversky, 1979). It posits that a decision-maker chooses in relation to an internal reference point – anything above it is considered a gain whereas anything below a loss. Its theoretical cousin is the energy budget rule from behavioral ecology (Kacelnik and Bateson, 1996), which describes the animal's risk attitude as dependent on its energy budget (positive or negative) while foraging. Constantinople et al. (2019b) showed that a reference-point model with lapse and history parameters fit their rats the best, thus, we cannot exclude the possibility that such model may be the best-fitting model for our rats overall.

Acknowledgments

We thank Yidi Chen, Anyu Fang, NengNeng Gao, Yingkun Li and CeQun Wang for technical assistance related to building and maintaining lab infrastructure as well as training animals and assisting with infusions. JCE acknowledges the support of the 111 project (Base B16018), the National Natural Science Foundation of China (NSFC) and the support of the NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai. EL and the human subjects research were supported by the National Science Foundation of China (NSFC-31750110461), International Young Scholar Grant and Shanghai Eastern Scholar Program (Shanghai Municipal Education Commission) Talents Young Scholar Grant.

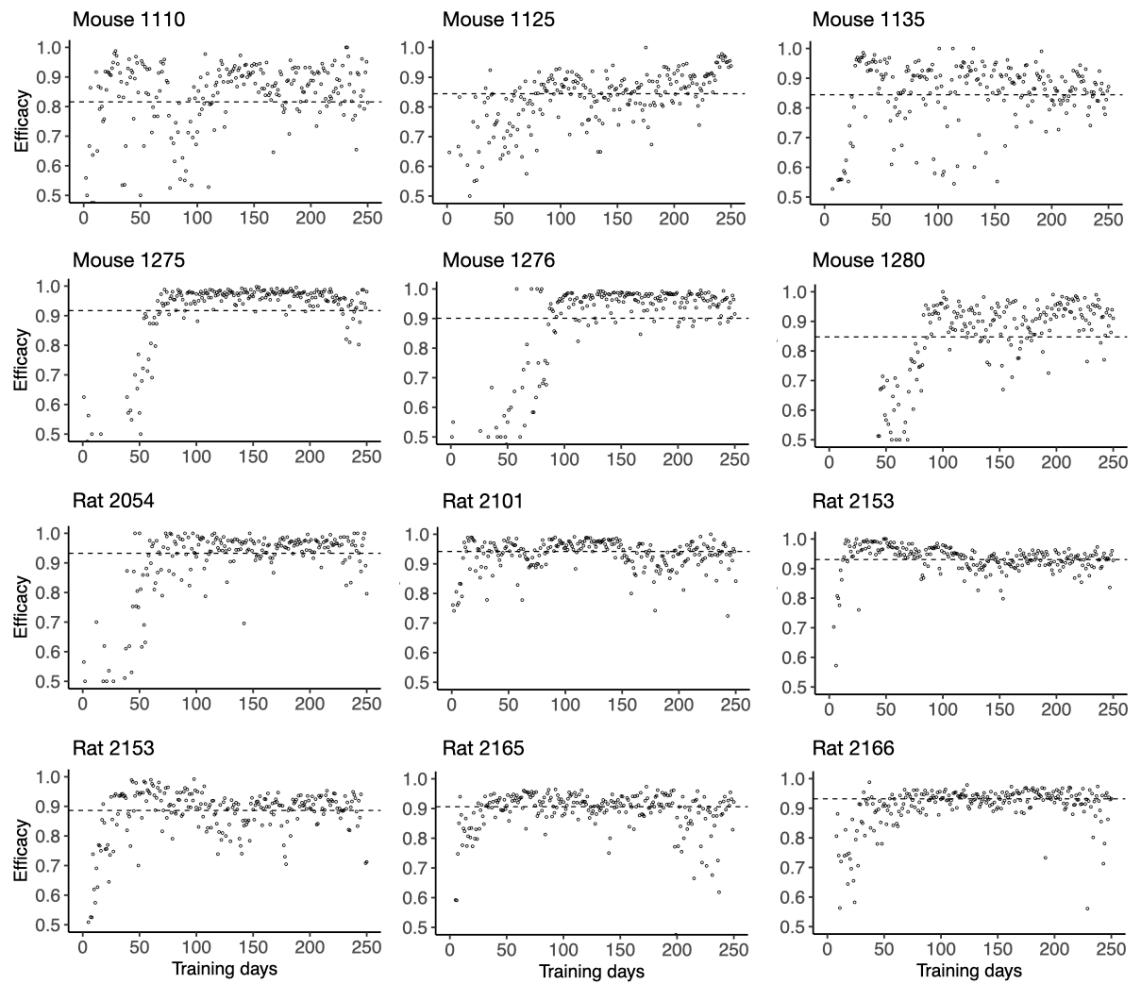


Figure S1. Example behavioral performance over training. Animal's 'efficacy' on each training session was computed by comparing its choices against that from a random agent and from a perfect expected value maximizer. See Methods for details. The dashed line indicates the median efficacy over 250 training days.

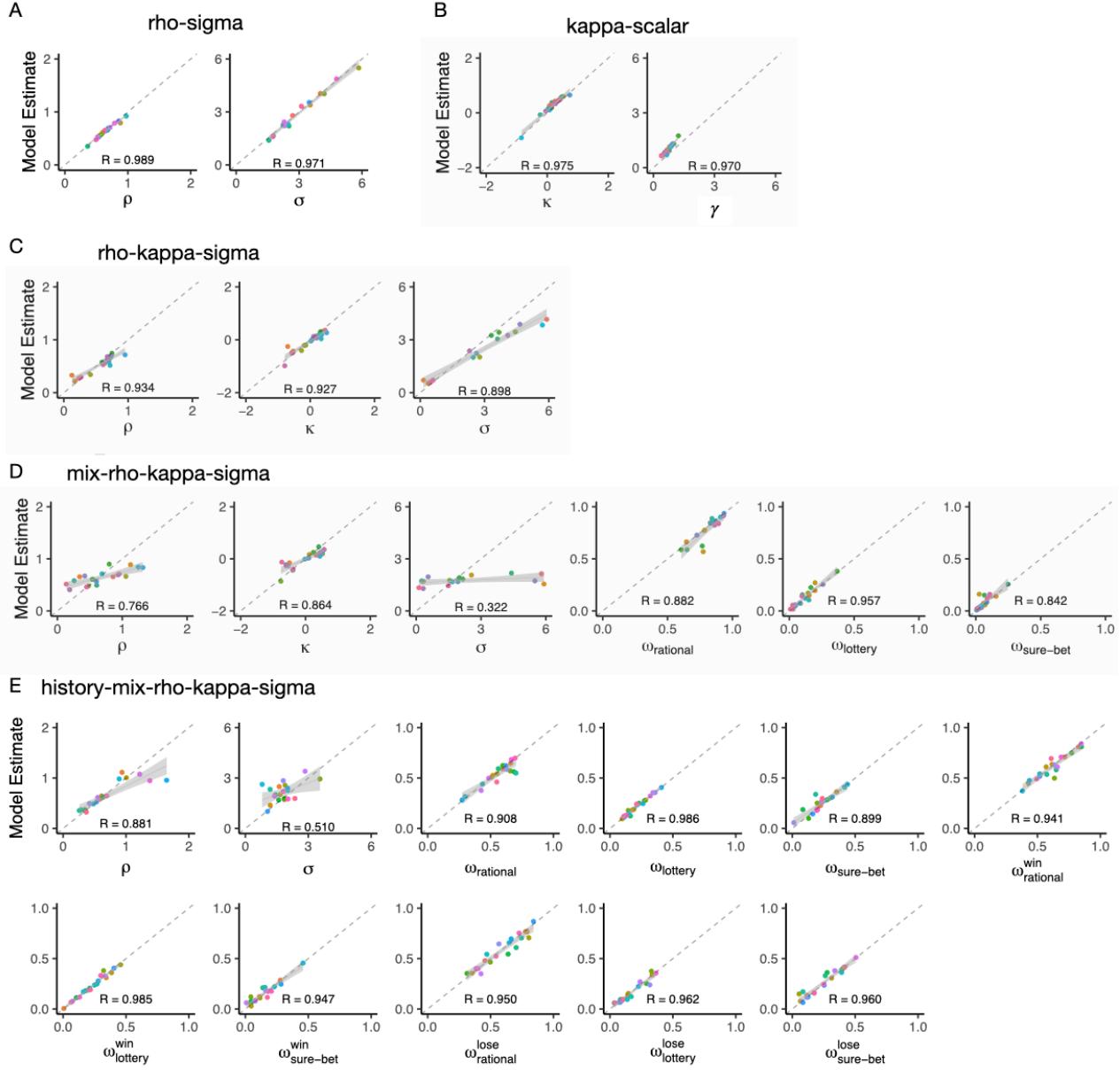


Figure S2. The models can recover the data-generating parameters well. Twenty Synthetic datasets were created by sampling from aforementioned prior distributions. The true parameter values are on the x-axis, MAP model estimates are on the y-axis. Color represents the identity of each synthetic dataset. Only 5 models are shown here, for a full list of sanity checks, see xxx.

Mouse

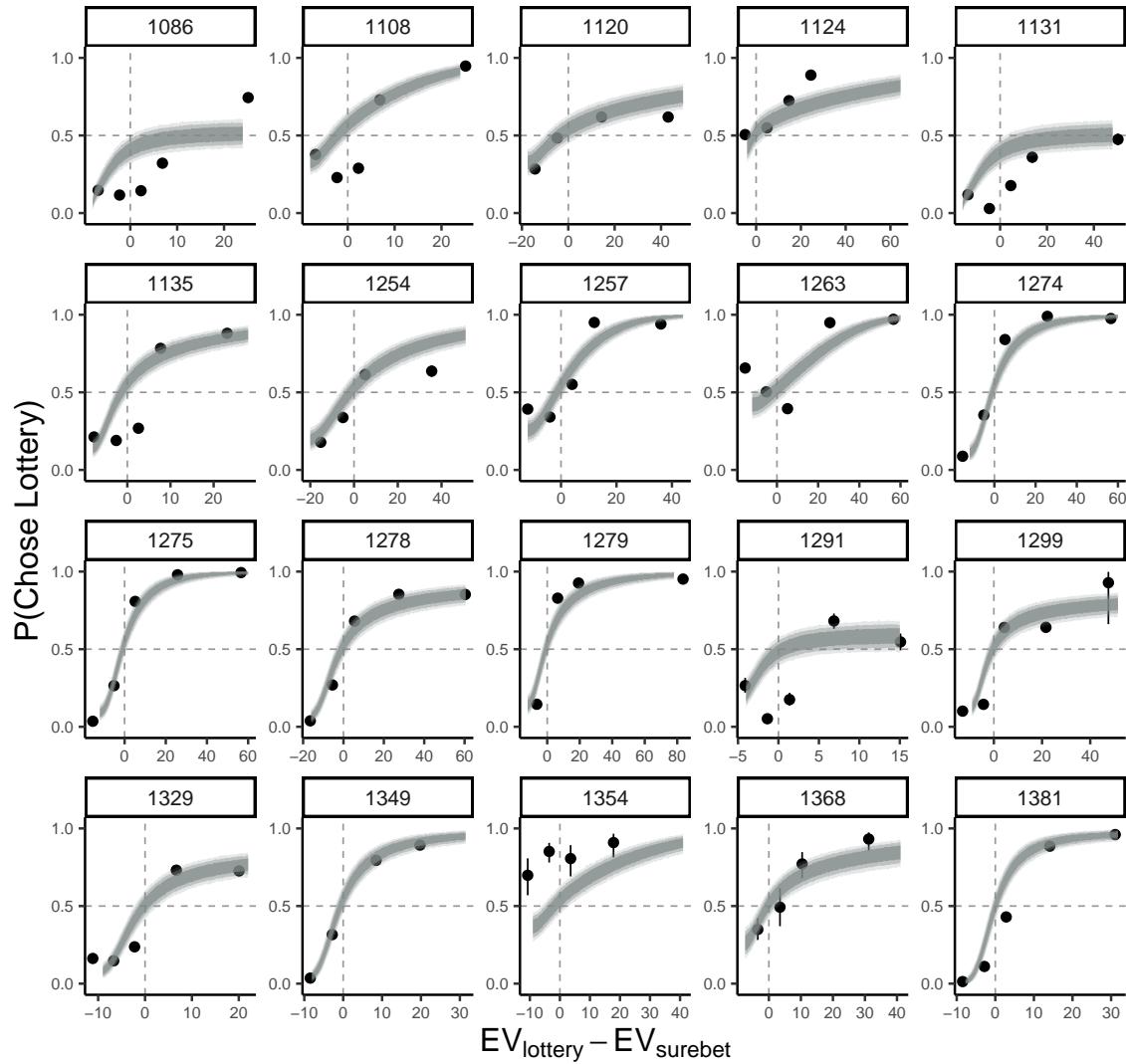


Figure S3. Example mouse's choices superimposed with predictions from the rho-sigma model. The circles with error bars are the binned mean and 95% binomial confidence intervals. The ribbons are model predictions generated using fitted parameters from the rho-sigma model. The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively.

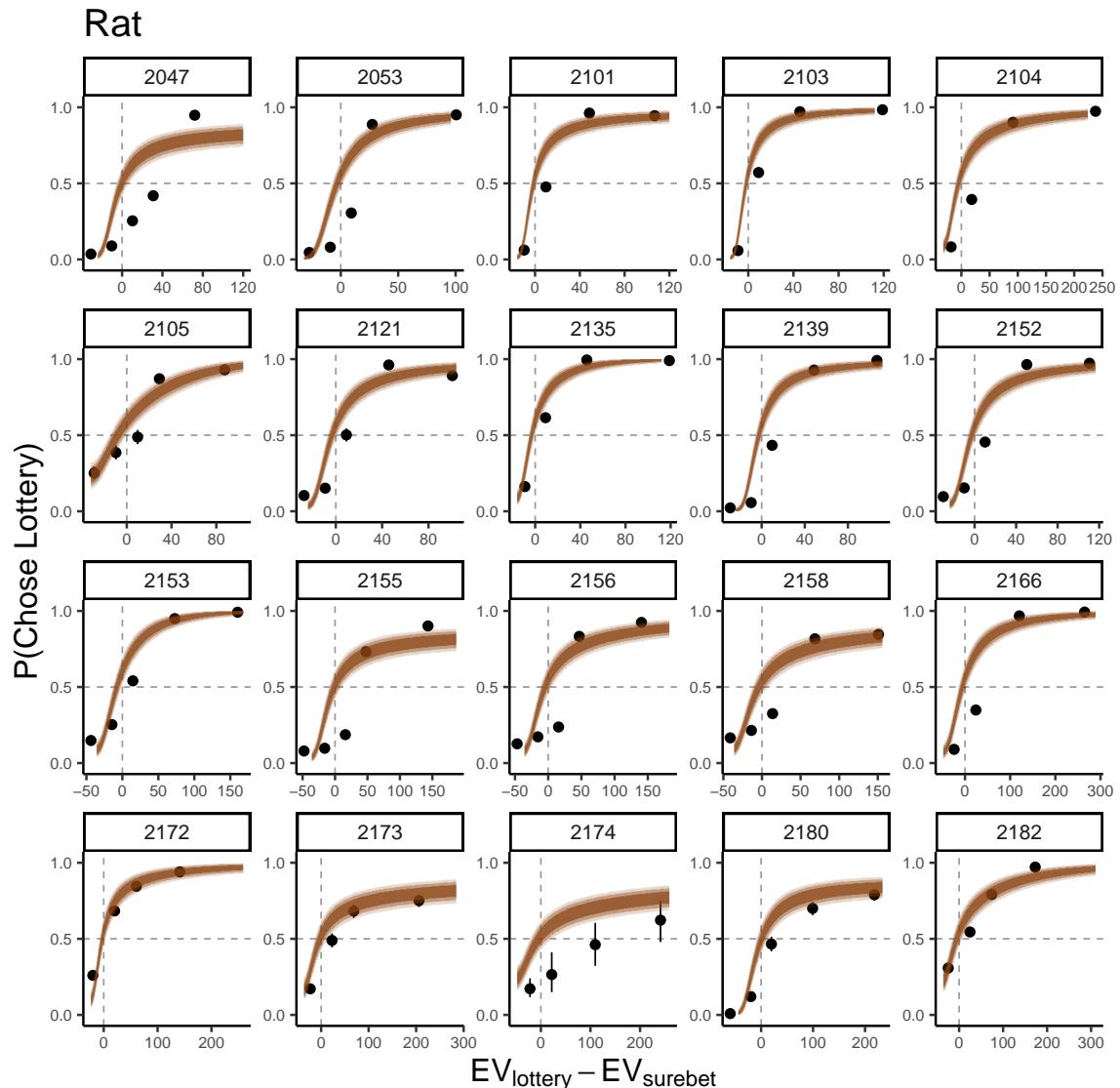


Figure S4. Example rat's choices superimposed with predictions from the rho-sigma model

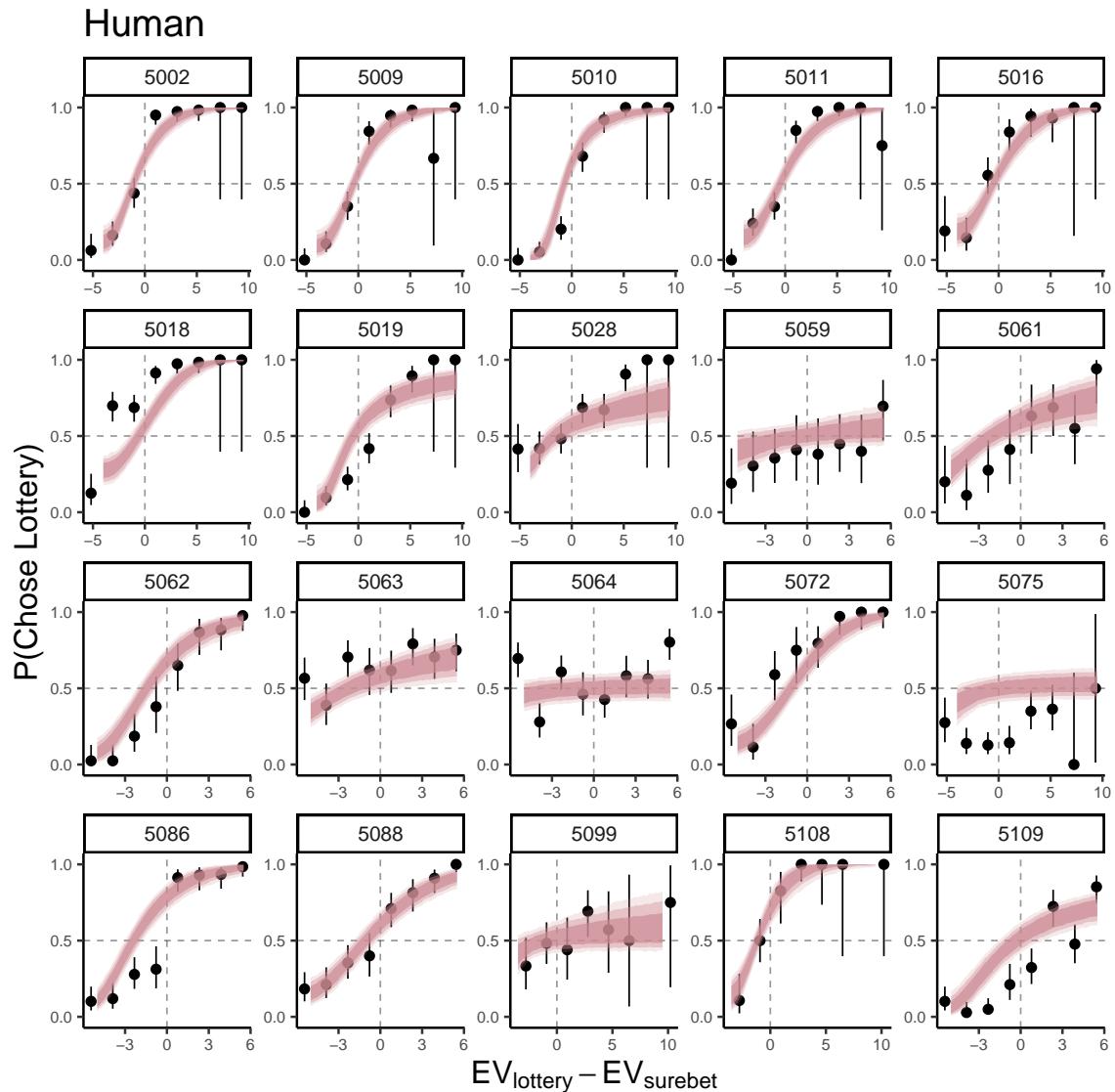


Figure S5. Example binned human's choices superimposed with predictions from the rho-sigma model.

Mouse

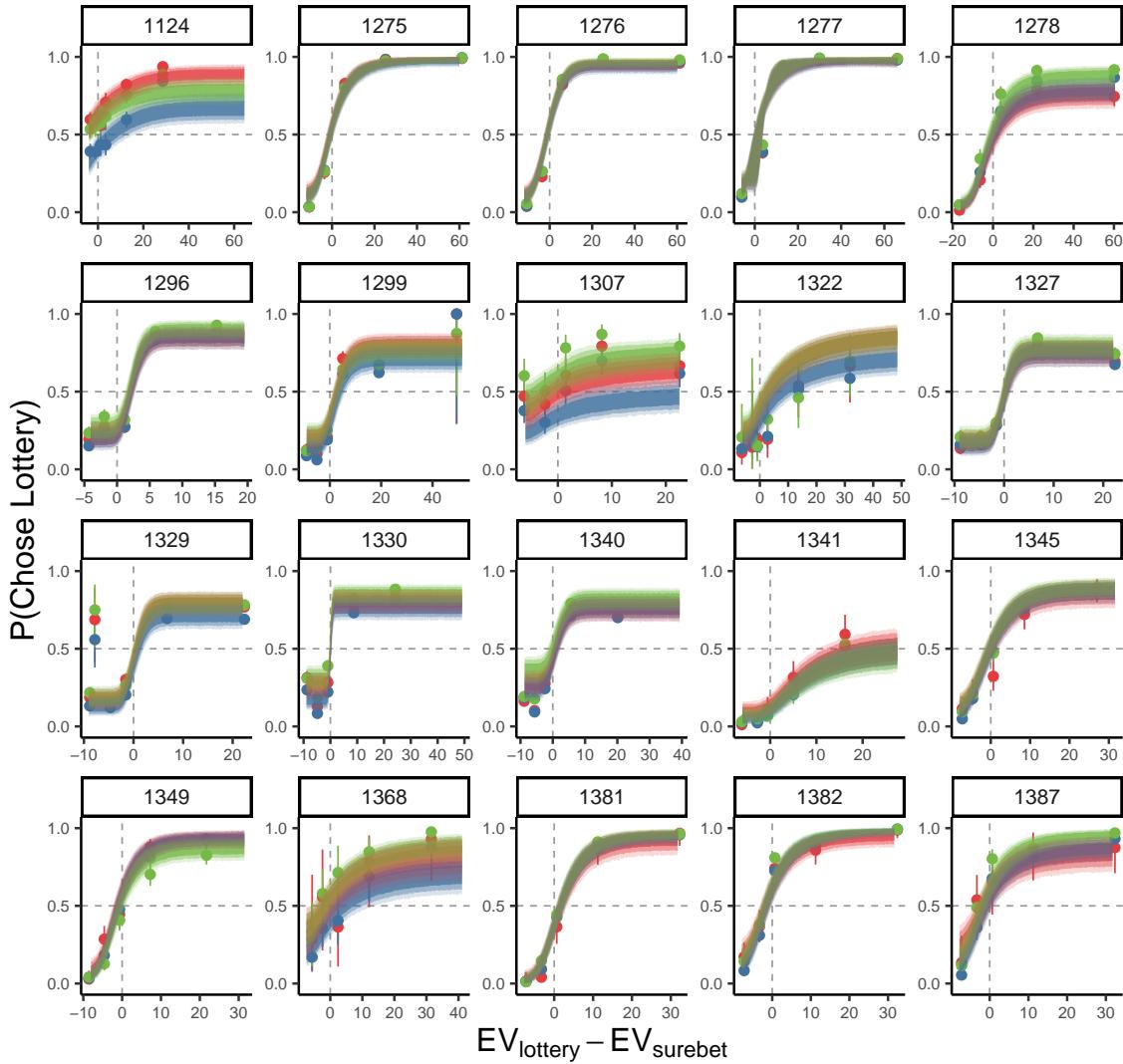


Figure S6. Example mouse's choices superimposed with predictions from the history-mix-rho-kappa-sigma model. The color represents the previous trial's outcome: lottery-win = green, lottery-lose = red, surebet = blue.

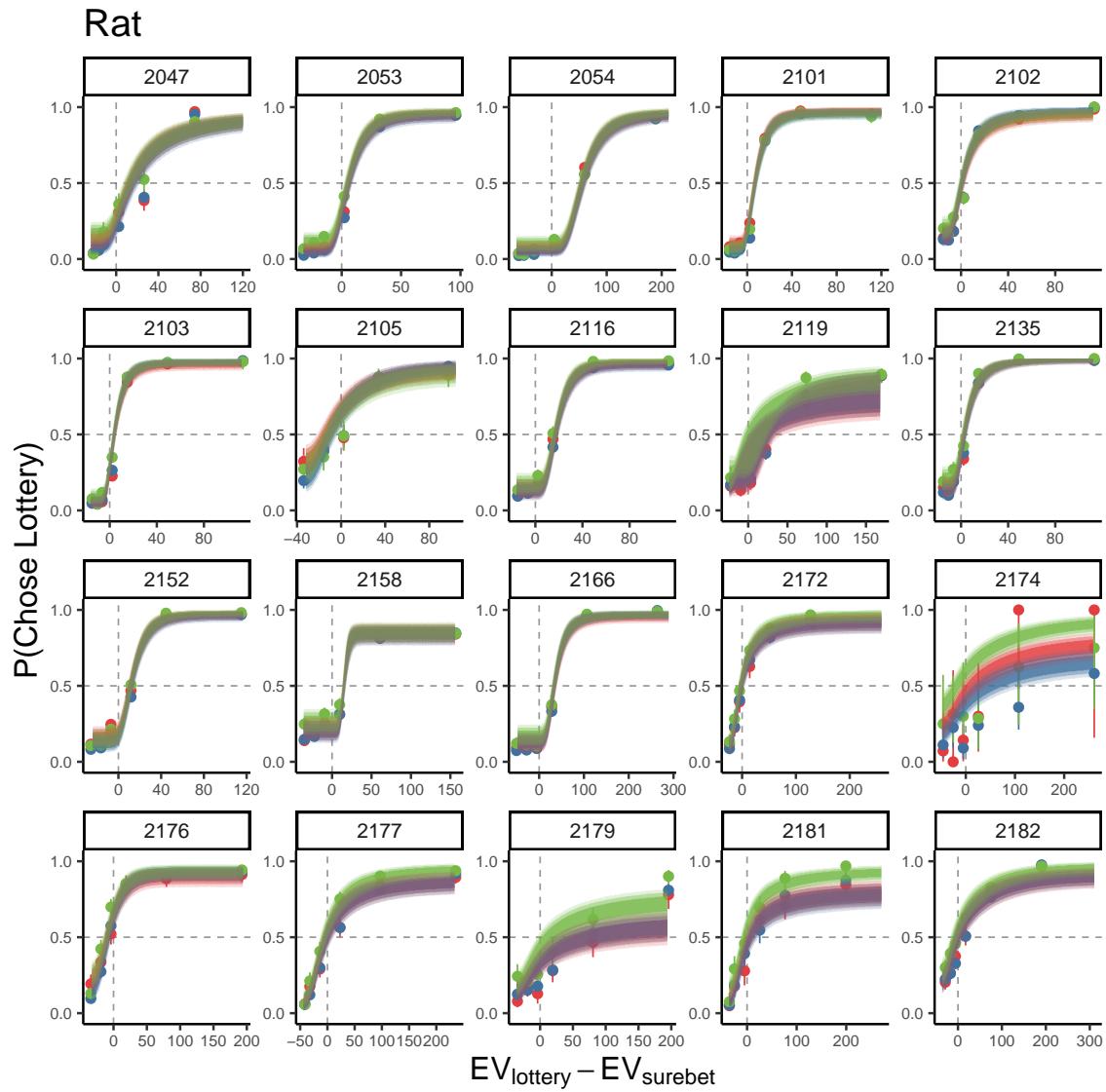


Figure S7. Example rat's choices superimposed with predictions from the history-mix-rho-kappa-sigma model.

Human

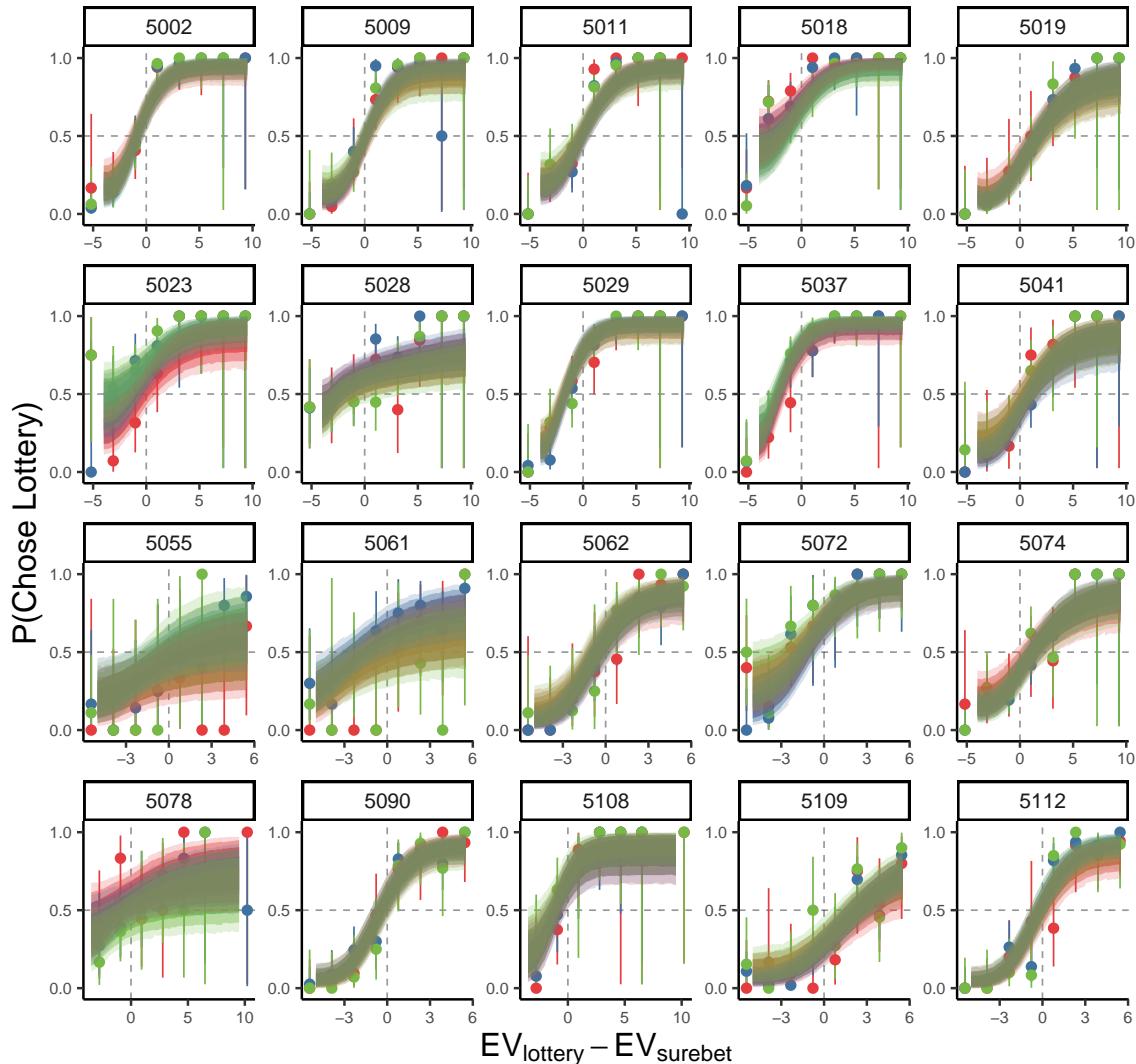


Figure S8. Example binned human's choices superimposed with predictions from the history-mix-rho-kappa-sigma model.

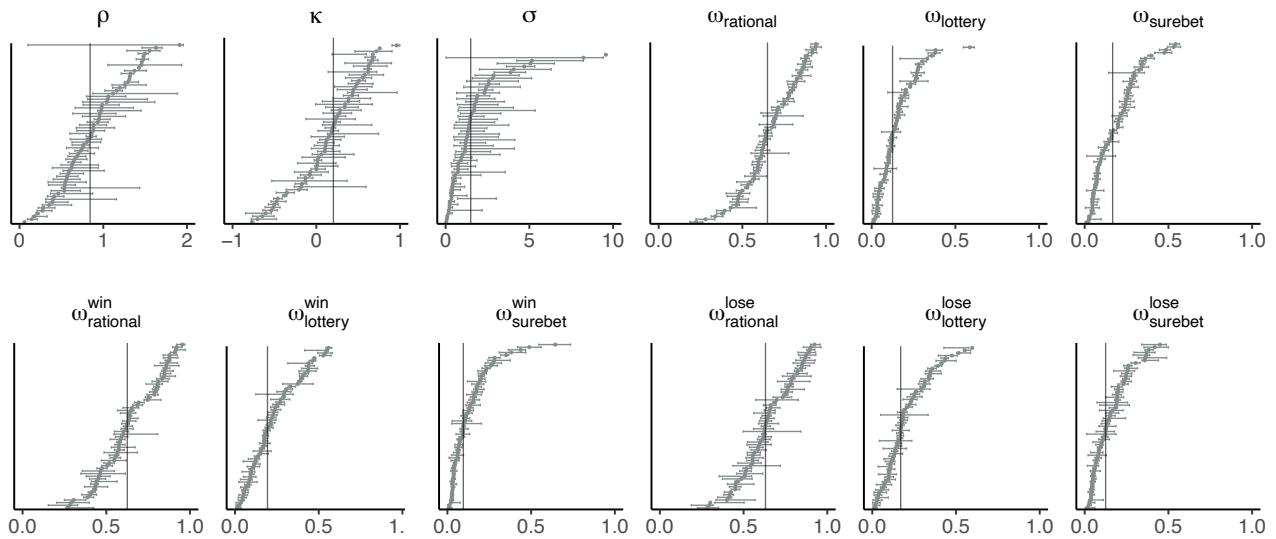


Figure S9. MAP parameter estimation with 95 % C.I. fit by the history-mixture-rho-kappa-sigma model for each mouse. The black bar indicates the median of concatenated posteriori samples across all subjects

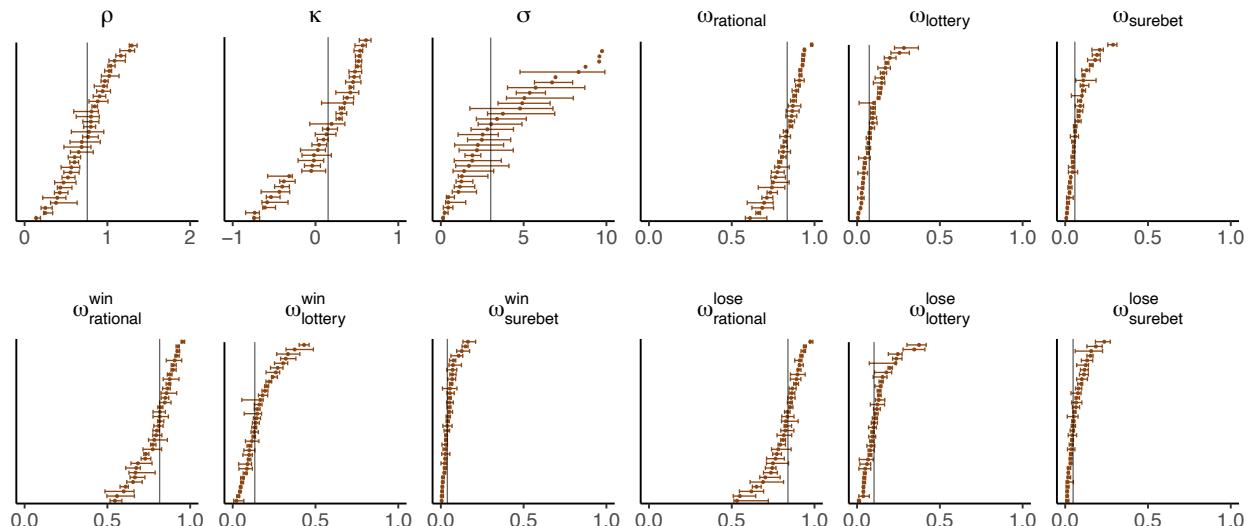


Figure S10. MAP parameter estimation with 95 % C.I. fit by the history-mixture-rho-kappa-sigma model for each rat.

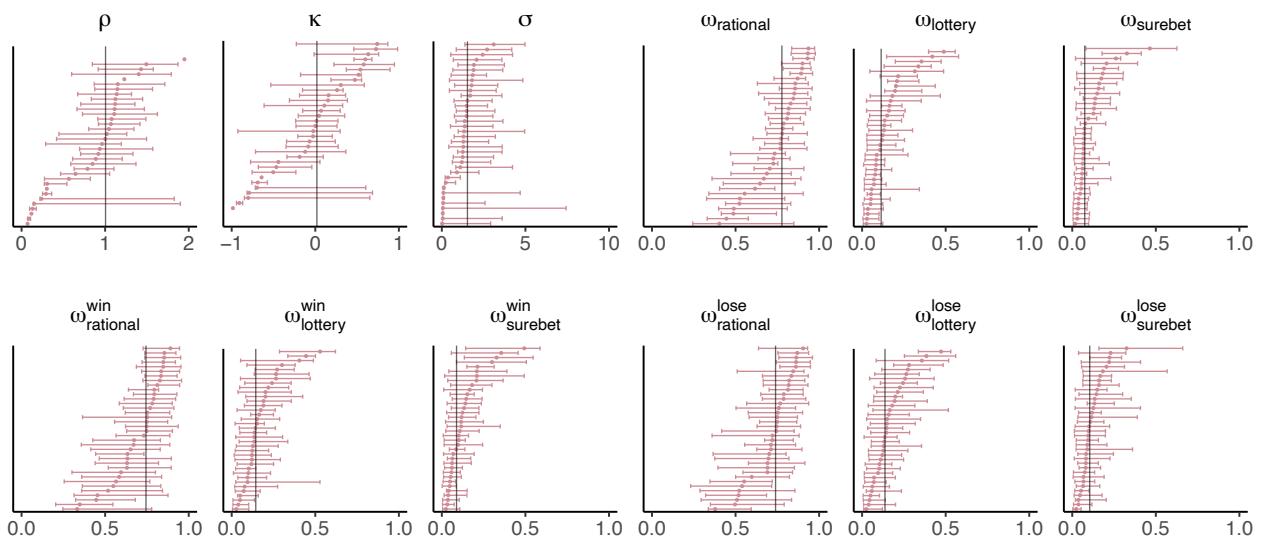


Figure S11. MAP parameter estimation with 95 % C.I. fit by the history-mixture-rho-kappa-sigma model for each human subject.

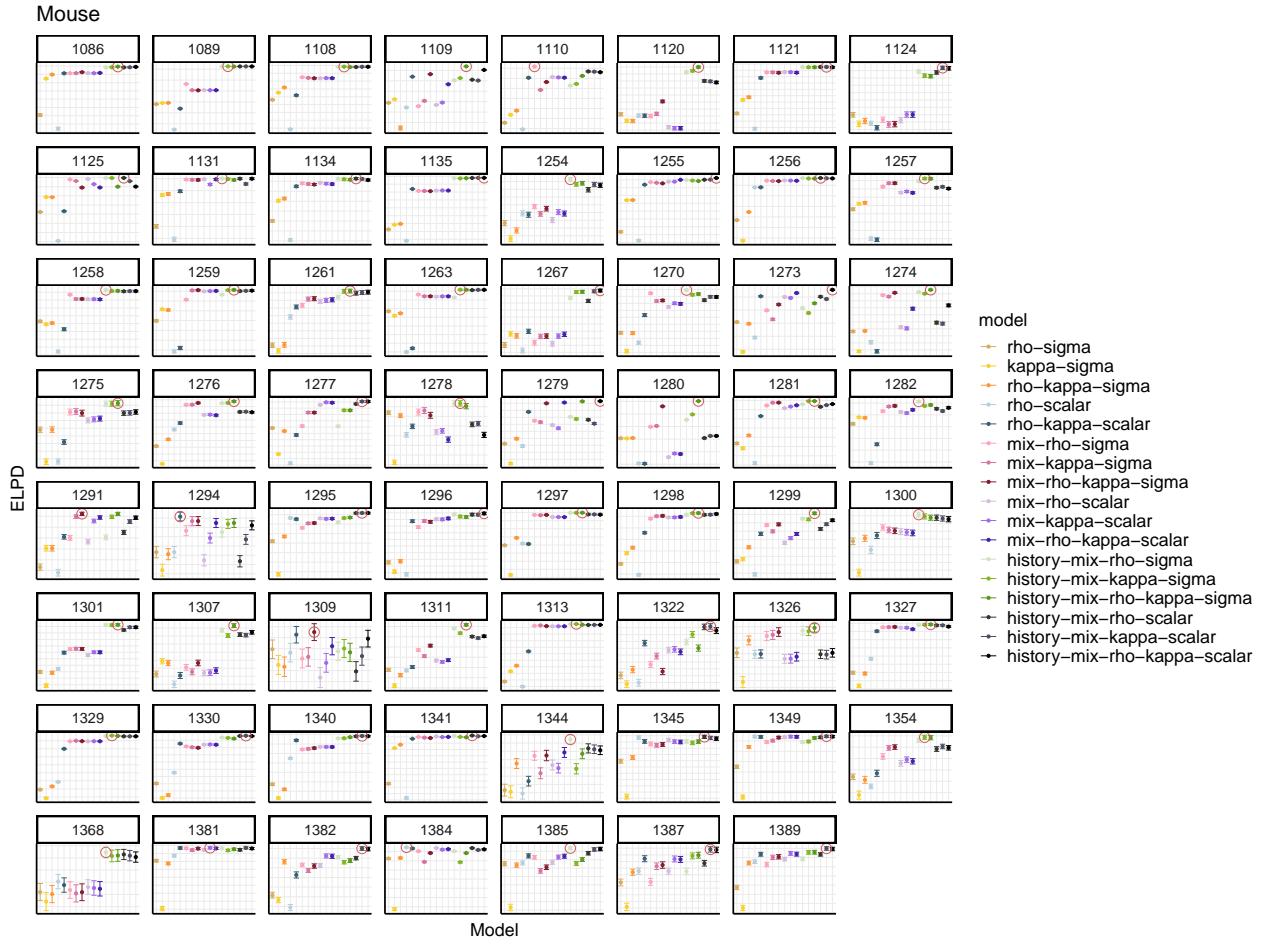


Figure S12. Ten-fold cross-validation results comparing all 18 models for the mouse population. The points with error bars are the expected log posterior density (ELPD) and its standard error on each animal's dataset, colored by model.

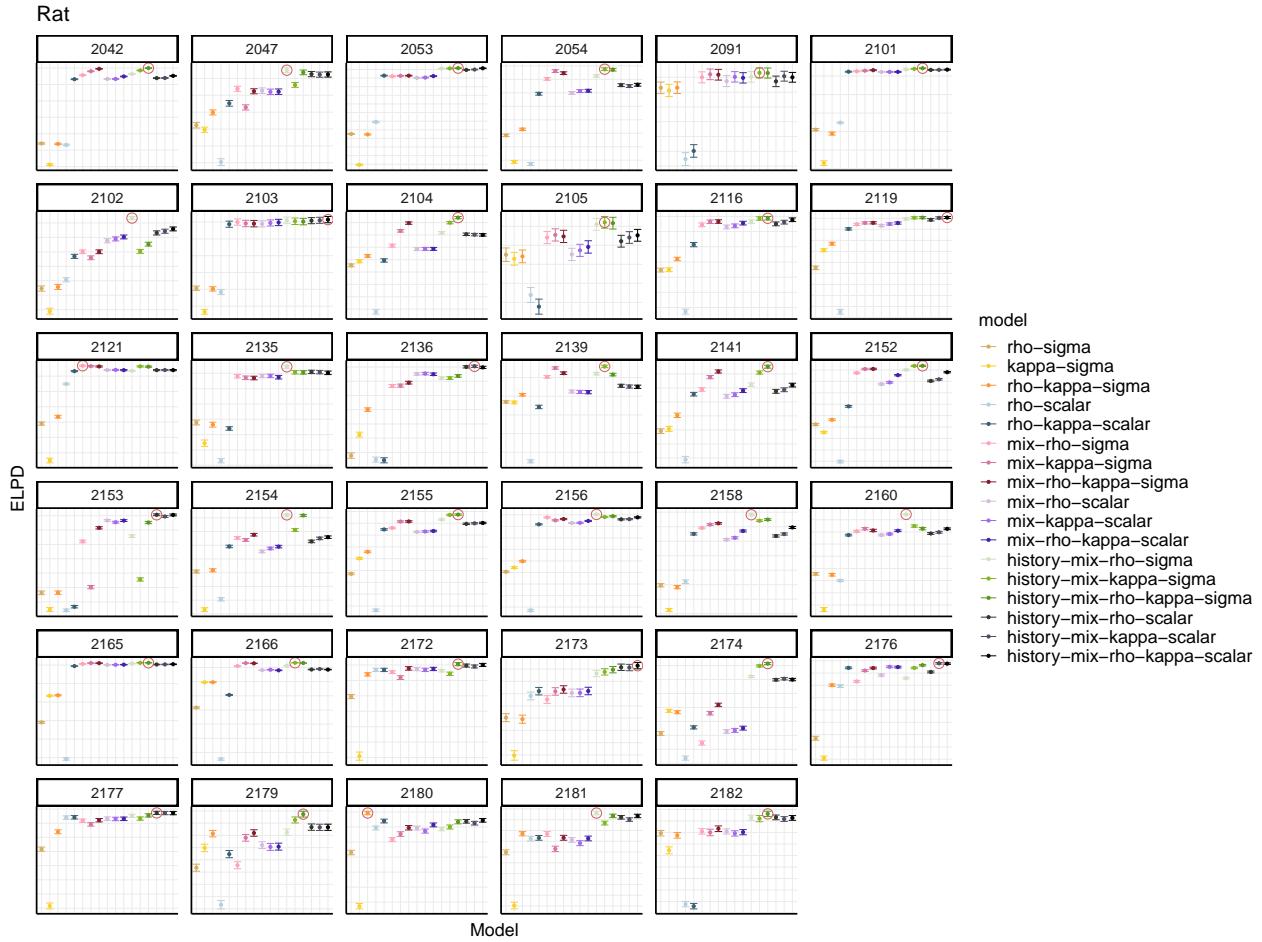


Figure S13. Ten-fold cross-validation results comparing all 18 models for the rat population. The points with error bars are the expected log posterior density (ELPD) and its standard error on each animal's dataset, colored by model.

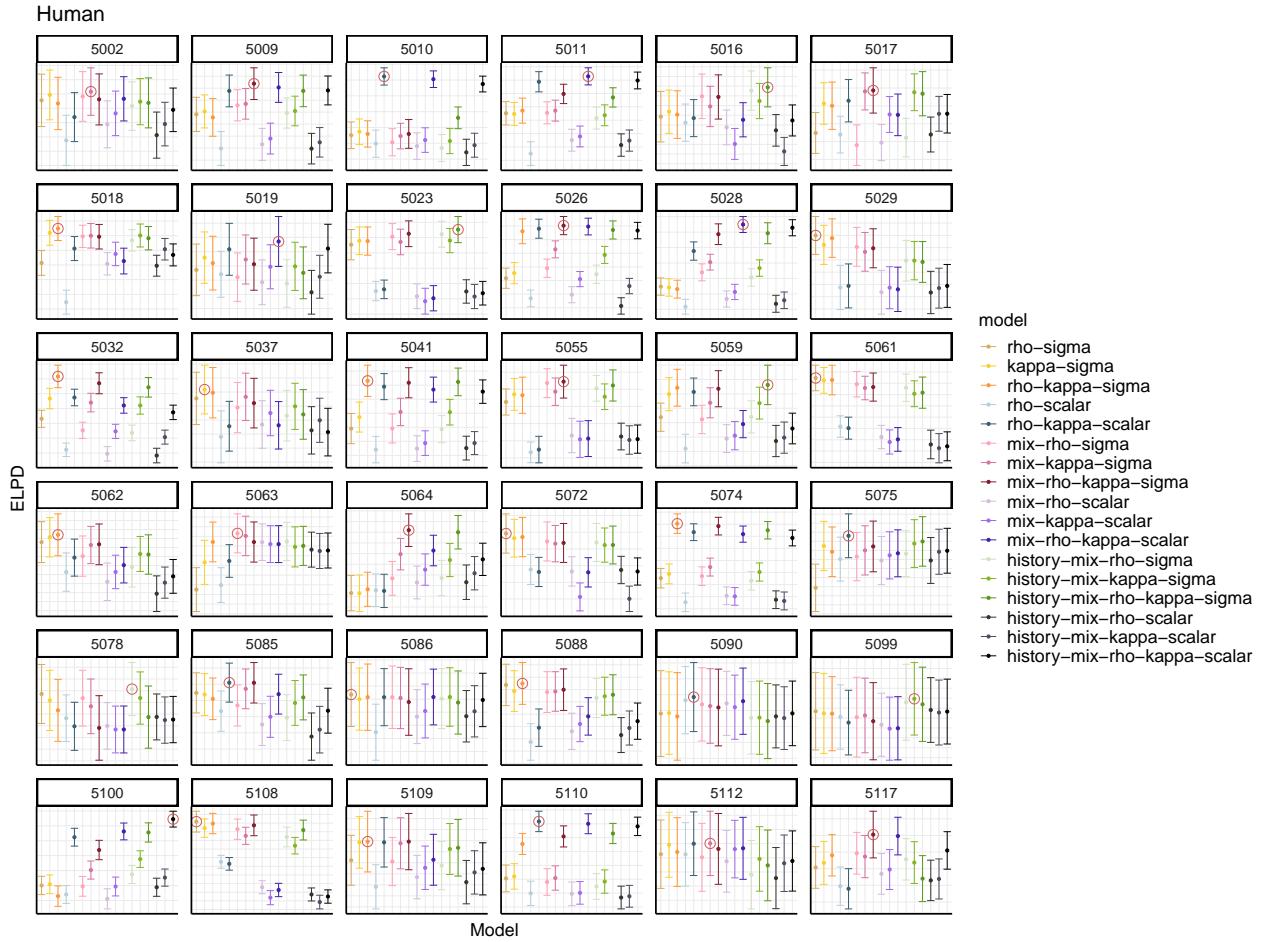


Figure S14. Ten-fold cross-validation results comparing all 18 models for the human population. The points with error bars are the expected log posterior density (ELPD) and its standard error on each animal's dataset, colored by model.

References

- 23and Me Research Team, eQTLgen Consortium, International Cannabis Consortium, Social Science Genetic Association Consortium, Richard Karlsson Linnér, Pietro Biroli, Edward Kong, S. Fleur W. Meddends, Robbee Wedow, Mark Alan Fontana, Maël Lebreton, Stephen P. Tino, Abdel Abdellaoui, Anke R. Hammerschlag, Michel G. Nivard, Aysu Okbay, Cornelius A. Rietveld, Pascal N. Timshel, Maciej Trzaskowski, Ronald de Vlaming, Christian L. Zünd, Yanchun Bao, Laura Budzugan, Ann H. Caplin, Chia-Yen Chen, Peter Eibich, Pierre Fontanillas, Juan R. Gonzalez, Peter K. Joshi, Ville Karhunen, Aaron Kleinman, Remy Z. Levin, Christina M. Lill, Gerardus A. Meddends, Gerard Muntané, Sandra Sanchez-Roige, Frank J. van Rooij, Erdogan Taskesen, Yang Wu, Futao Zhang, Adam Auton, Jason D. Boardman, David W. Clark, Andrew Conlin, Conor C. Dolan, Urs Fischbacher, Patrick J. F. Groenen, Kathleen Mullan Harris, Gregor Hasler, Albert Hofman, Mohammad A. Ikram, Sonia Jain, Robert Karlsson, Ronald C. Kessler, Maarten Kooyman, James MacKillop, Minna Männikkö, Carlos Morcillo-Suarez, Matthew B. McQueen, Klaus M. Schmidt, Melissa C. Smart, Matthias Sutter, A. Roy Thurik, André G. Uitterlinden, Jon White, Harriet de Wit, Jian Yang, Lars Bertram, Dorret I. Boomsma, Tõnu Esko, Ernst Fehr, David A. Hinds, Magnus Johannesson, Meena Kumari, David Laibson, Patrik K. E. Magnusson, Michelle N. Meyer, Arcadi Navarro, Abraham A. Palmer, Tune H. Pers, Danielle Posthuma, Daniel Schunk, Murray B. Stein, Rauli Svento, Henning Tiemeier, Paul R. H. J. Timmers, Patrick Turley, Robert J. Ursano, Gert G. Wagner, James F. Wilson, Jacob Gratten, James J. Lee, David Cesarini, Daniel J. Benjamin, Philipp D. Koellinger, and Jonathan P. Beauchamp. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature Genetics*, 51(2):245–257, February 2019. ISSN 1061-4036, 1546-1718. doi: 10.1038/s41588-018-0309-3.
- Arman Abrahamyan, Laura Luz Silva, Steven C. Dakin, Matteo Carandini, and Justin L. Gardner. Adaptable history biases in human perceptual decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 113(25):E3548–3557, June 2016. ISSN 1091-6490. doi: 10.1073/pnas.1518786113.
- Serge H. Ahmed. Individual decision-making in the causal pathway to addiction: Contributions and limitations of rodent models. *Pharmacology Biochemistry and Behavior*, 164:22–31, January 2018. ISSN 0091-3057. doi: 10.1016/j.pbb.2017.07.005.
- Athena Akrami, Charles D. Kopec, Mathew E. Diamond, and Carlos D. Brody. Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature*, 554(7692):368–372, February 2018. ISSN 1476-4687. doi: 10.1038/nature25510.
- Carsten Allefeld, Chun Siong Soon, Carsten Bogler, Jakob Heinze, and John-Dylan Haynes. Sequential dependencies between trials in free choice tasks. *arXiv:1311.0753 [q-bio]*, November 2013.
- Andrey P. Anokhin, Simon Golosheykin, Julia Grant, and Andrew C. Heath. Heritability of Risk-Taking in Adolescence: A Longitudinal Twin Study. *Twin Research and Human Genetics*, 12(4):366–371, August 2009. ISSN 1832-4274, 1839-2628. doi: 10.1375/twin.12.4.366.
- Fuat Balci, David Freestone, and Charles R. Gallistel. Risk assessment in man and mouse. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7):2459–2463, February 2009. ISSN 1091-6490. doi: 10.1073/pnas.0812709106.
- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1):1–48, October 2015. ISSN 1548-7660. doi: 10.18637/jss.v067.i01.
- Eduard Brandstätter, Gerd Gigerenzer, and Ralph Hertwig. The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113(2):409–432, 2006. ISSN 1939-1471(Electronic),0033-295X(Print). doi: 10.1037/0033-295X.113.2.409.
- Matteo Carandini and Anne K. Churchland. Probing perceptual decisions in rodents. *Nature Neuroscience*, 16(7):824–831, July 2013. ISSN 1097-6256. doi: 10.1038/nn.3410.

-
- Bob Carpenter, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 20:1–37, 2016.
- Vikram S. Chib, Antonio Rangel, Shinsuke Shimojo, and John P. O'Doherty. Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *Journal of Neuroscience*, 29(39):12315–12320, September 2009. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2575-09.2009.
- Emma A. D. Clifton, John R. B. Perry, Fumiaki Immura, Luca A. Lotta, Soren Brage, Nita G. Forouhi, Simon J. Griffin, Nicholas J. Wareham, Ken K. Ong, and Felix R. Day. Genome-wide association study for risk taking propensity indicates shared pathways with body mass index. *Communications Biology*, 1(1):36, December 2018. ISSN 2399-3642. doi: 10.1038/s42003-018-0042-6.
- Christine M Constantinople, Alex T Piet, Peter Bibawi, Athena Akrami, Charles Kopec, and Carlos D Brody. Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *eLife*, 8:e49744, November 2019a. ISSN 2050-084X. doi: 10.7554/eLife.49744.
- Christine M. Constantinople, Alex T. Piet, and Carlos D. Brody. An Analysis of Decision under Risk in Rats. *Current Biology*, 29(12):2066–2074.e5, June 2019b. ISSN 09609822. doi: 10.1016/j.cub.2019.05.013.
- Peter Dayan and Angela J Yu. Expected and Unexpected Uncertainty: ACh and NE in the Neocortex. In *Advances in Neural Information Processing Systems*, volume 15. MIT Press, 2003.
- J. W. de Gee, K. Tsetsos, L. Schwabe, A.E. Urai, D. A. McCormick, M. J. McGinley, and T. H. Donner. Phasic arousal suppresses biases in mice and humans across domains of decision-making. Preprint, Neuroscience, October 2018.
- Micheal L. Dent, Laurel A. Screven, and Anastasiya Kobraina. Hearing in Rodents. In Micheal L. Dent, Richard R. Fay, and Arthur N. Popper, editors, *Rodent Bioacoustics*, Springer Handbook of Auditory Research, pages 71–105. Springer International Publishing, Cham, 2018. ISBN 978-3-319-92495-3. doi: 10.1007/978-3-319-92495-3-4.
- Catherine C. Eckel and Philip J. Grossman. Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, 23(4):281–295, July 2002. ISSN 1090-5138. doi: 10.1016/S1090-5138(02)00097-1.
- B. R Eisenreich and Benjamin Y. Hayden. Risk aversion in macaques in a freely moving patch-leaving foraging task. Preprint, Animal Behavior and Cognition, October 2018.
- Jeffrey C Erlich, Bingni W Brunton, Chunyu A Duan, Timothy D Hanks, and Carlos D Brody. Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 4, 2015. doi: 10.7554/eLife.05457.
- Shiva Farashahi, Habiba Azab, Benjamin Hayden, and Alireza Soltani. On the Flexibility of Basic Risk Attitudes in Monkeys. *The Journal of Neuroscience*, 38(18):4383–4398, May 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2260-17.2018.
- Shiva Farashahi, Christopher H. Donahue, Benjamin Y. Hayden, Daeyeol Lee, and Alireza Soltani. Flexible combination of reward information across primates. *Nature Human Behaviour*, 3(11):1215–1224, November 2019. ISSN 2397-3374. doi: 10.1038/s41562-019-0714-3.
- Stan B. Floresco, David R. Montes, Maric M.T. Tse, and Mieke van Holstein. Differential contributions of nucleus accumbens subregions to cue-guided risk/reward decision making and implementation of conditional rules. *The Journal of Neuroscience*, pages 3191–17, January 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3191-17.2018.

-
- Joshua I. Gold, Chi-Tat Law, Patrick Connolly, and Sharath Bennur. The Relative Influences of Priors and Sensory Evidence on an Oculomotor Decision Variable During Perceptual Learning. *Journal of Neurophysiology*, 100(5):2653–2668, November 2008. ISSN 0022-3077. doi: 10.1152/jn.90629.2008.
- Timothy D. Hanks and Christopher Summerfield. Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93(1):15–31, January 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.12.003.
- Todd A. Hare, John O’Doherty, Colin F. Camerer, Wolfram Schultz, and Antonio Rangel. Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors. *Journal of Neuroscience*, 28(22):5623–5630, May 2008. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1309-08.2008.
- Joop Hartog, Ada Ferrer-i-Carbonell, and Nicole Jonker. Linking Measured Risk Aversion to Individual Characteristics. *Kyklos*, 55(1):3–26, 2002. ISSN 1467-6435. doi: 10.1111/1467-6435.00175.
- Johannes Haushofer and Ernst Fehr. On the psychology of poverty. *Science*, 344(6186):862–867, May 2014. doi: 10.1126/science.1232491.
- Benjamin Y. Hayden and Michael L. Platt. Temporal Discounting Predicts Risk Sensitivity in Rhesus Macaques. *Current Biology*, 17(1):49–53, January 2007. ISSN 0960-9822. doi: 10.1016/j.cub.2006.10.055.
- Sarah Heilbronner, Benjamin Hayden, and Michael Platt. Decision Salience Signals in Posterior Cingulate Cortex. *Frontiers in Neuroscience*, 5:55, 2011. ISSN 1662-453X. doi: 10.3389/fnins.2011.00055.
- Sarah R Heilbronner. Modeling risky decision-making in nonhuman animals: Shared core features. *Current Opinion in Behavioral Sciences*, 16:23–29, August 2017. ISSN 2352-1546. doi: 10.1016/j.cobeha.2017.03.001.
- Sarah R. Heilbronner and Benjamin Y. Hayden. Contextual Factors Explain Risk-Seeking Preferences in Rhesus Monkeys. *Frontiers in Neuroscience*, 7, 2013. ISSN 1662-4548. doi: 10.3389/fnins.2013.00007.
- Sarah R. Heilbronner and Benjamin Y. Hayden. The description-experience gap in risky choice in nonhuman primates. *Psychonomic Bulletin & Review*, 23(2):593–600, April 2016. ISSN 1531-5320. doi: 10.3758/s13423-015-0924-2.
- Ralph Hertwig and Ido Erev. The description-experience gap in risky choice. *Trends in Cognitive Sciences*, 13 (12):517–523, December 2009. ISSN 1364-6613. doi: 10.1016/j.tics.2009.09.004.
- Ralph Hertwig, Greg Barron, Elke U. Weber, and Ido Erev. Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, 15(8):534–539, August 2004. ISSN 0956-7976, 1467-9280. doi: 10.1111/j.0956-7976.2004.00715.x.
- Tetiana Hill, Petko Kusev, and Paul van Schaik. Choice Under Risk: How Occupation Influences Preferences. *Frontiers in Psychology*, 10:2003, 2019. ISSN 1664-1078. doi: 10.3389/fpsyg.2019.02003.
- Charles A Holt, Susan K Laury, et al. Risk aversion and incentive effects. *American economic review*, 92(5): 1644–1655, 2002.
- Eun Jung Hwang, Jeffrey E. Dahlen, Madan Mukundan, and Takaki Komiyama. History-based action selection bias in posterior parietal cortex. *Nature Communications*, 8(1):1242, November 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-01356-z.
- Alex Kacelnik and Melissa Bateson. Risky Theories—The Effects of Variance on Foraging Decisions. *Integrative and Comparative Biology*, 36(4):402–434, September 1996. ISSN 1540-7063. doi: 10.1093/icb/36.4.402.
- Alex Kacelnik and Fausto Brito e Abreu. Risky Choice and Weber’s Law. *Journal of Theoretical Biology*, 194 (2):289–298, September 1998. ISSN 0022-5193. doi: 10.1006/jtbi.1998.0763.

-
- Daniel Kahneman and Shane Frederick. Representativeness Revisited: Attribute Substitution in Intuitive Judgment. *Heuristics and biases: The psychology of intuitive judgment*, 49:49–81, July 2002. ISSN 9780521792608. doi: 10.1017/CBO9780511808098.004.
- Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263, March 1979. ISSN 00129682. doi: 10.2307/1914185.
- Alexander Klos, Elke U. Weber, and Martin Weber. Investment Decisions and Time Horizon: Risk Perception and Risk Behavior in Repeated Gambles. *Management Science*, 51(12):1777–1790, December 2005. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.1050.0429.
- Frank Hyneman Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin, 1921.
- Jerald D. Kralik, Eric R. Xu, Emily J. Knight, Sara A. Khan, and William J. Levine. When Less Is More: Evolutionary Origins of the Affect Heuristic. *PLOS ONE*, 7(10):e46240, October 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0046240.
- Joshua D. Larkin, Nicole L. Jenni, and Stan B. Floresco. Modulation of risk/reward decision making by dopaminergic transmission within the basolateral amygdala. *Psychopharmacology*, 233(1):121–136, 2016.
- Irwin P. Levin, Mary A. Snyder, and Daniel P. Chapman. The Interaction of Experiential and Situational Factors and Gender in a Simulated Risky Decision-Making Task. *The Journal of Psychology*, 122(2):173–181, March 1988. ISSN 0022-3980. doi: 10.1080/00223980.1988.9712703.
- Dino J Levy and Paul W Glimcher. The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, 22(6):1027–1038, December 2012. ISSN 09594388. doi: 10.1016/j.conb.2012.06.001.
- Lola L. Lopes. Decision making in the short run. *Journal of Experimental Psychology: Human Learning & Memory*, 7(5):377–385, 1981. ISSN 0096-1515. doi: 10.1037/0278-7393.7.5.377.
- Elliot A. Ludvig and Marcia L. Spetch. Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? *PLOS ONE*, 6(6):e20262, June 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0020262.
- Evgeniya Lukinova, Yuyue Wang, Steven F Lehrer, and Jeffrey C Erlich. Time preferences are reliable across time-horizons and verbal versus experiential tasks. *eLife*, 8:e39656, February 2019. ISSN 2050-084X. doi: 10.7554/eLife.39656.
- Harry Markowitz. The Utility of Wealth. *Journal of Political Economy*, 60(2):151–158, April 1952. ISSN 0022-3808, 1537-534X. doi: 10.1086/257177.
- Harry M. Markowitz. *Portfolio Selection*. Yale University Press, January 1968. ISBN 978-0-300-19167-7.
- Barnaby Marsh and Alex Kacelnik. Framing effects and risky decisions in starlings. *Proceedings of the National Academy of Sciences*, 99(5):3352–3355, March 2002. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.042491999.
- Andrew T. Marshall and Kimberly Kirkpatrick. The effects of the previous outcome on probabilistic choice in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 39(1):24, 2013. ISSN 1939-2184. doi: 10.1037/a0030765.
- Allison N. McCoy and Michael L. Platt. Expectations and outcomes: Decision-making in the primate brain. *Journal of Comparative Physiology A*, 191(3):201–211, March 2005. ISSN 0340-7594, 1432-1351. doi: 10.1007/s00359-004-0565-9.
- Ari S Morcos and Christopher D Harvey. History-dependent variability in population dynamics during evidence accumulation in cortex. *Nature Neuroscience*, 19(12):1672–1681, December 2016. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4403.

-
- Hansjörg Neth and Gerd Gigerenzer. Heuristics: Tools for an Uncertain World. page 18.
- Nader Nikbakht, Azadeh Tafreshiha, Davide Zoccolan, and Mathew E Diamond. Supralinear and supramodal integration of visual and tactile signals in rats: Psychophysics and neuronal mechanisms. *Neuron*, 97(3): 626–639, 2018.
- Yael Niv, Jeffrey A. Edlund, Peter Dayan, and John P. O'Doherty. Neural Prediction Errors Reveal a Risk-Sensitive Reinforcement-Learning Process in the Human Brain. *Journal of Neuroscience*, 32(2): 551–562, January 2012. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.5498-10.2012.
- Martin O'Neill and Wolfram Schultz. Coding of Reward Risk by Orbitofrontal Neurons Is Mostly Distinct from Coding of Reward Value. *Neuron*, 68(4):789–800, November 2010. ISSN 0896-6273. doi: 10.1016/j.neuron.2010.09.031.
- Thorsten Pachur, Ralph Hertwig, Gerd Gigerenzer, and Eduard Brandstätter. Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology*, 4:646, 2013. ISSN 1664-1078. doi: 10.3389/fpsyg.2013.00646.
- Camillo Padoa-Schioppa. Neuronal origins of choice variability in economic decisions. *Neuron*, 80(5): 1322–1336, December 2013. ISSN 1097-4199. doi: 10.1016/j.neuron.2013.09.013.
- Elise Payzan-LeNestour, Simon Dunne, Peter Bossaerts, and John P. O'Doherty. The Neural Representation of Unexpected Uncertainty during Value-Based Decision Making. *Neuron*, 79(1):191–201, July 2013. ISSN 0896-6273. doi: 10.1016/j.neuron.2013.04.037.
- Sashank Pisupati, Lital Chartarifsky-Lynn, Anup Khanal, and Anne K Churchland. Lapses in perceptual decisions reflect exploration. *eLife*, 10:e55490, January 2021. ISSN 2050-084X. doi: 10.7554/eLife.55490.
- Li-Lin Rao, Yuan Zhou, Dang Zheng, Liu-Qing Yang, and Shu Li. Genetic Contribution to Variation in Risk Taking: A Functional MRI Twin Study of the Balloon Analogue Risk Task. *Psychological Science*, 29(10): 1679–1691, October 2018. ISSN 0956-7976, 1467-9280. doi: 10.1177/0956797618779961.
- Robb B. Rutledge, Peter Smittenaar, Peter Zeidman, Harriet R. Brown, Rick A. Adams, Ulman Lindenberger, Peter Dayan, and Raymond J. Dolan. Risk Taking for Potential Reward Decreases across the Lifespan. *Current Biology*, 26(12):1634–1639, June 2016. ISSN 0960-9822. doi: 10.1016/j.cub.2016.05.017.
- Paul Samuelson. Risk and uncertainty: A fallacy of large numbers. *Scientia*, 57(98), 1963.
- Guillaume Sescousse, Xavier Caldú, Bàrbara Segura, and Jean-Claude Dreher. Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 37(4):681–696, May 2013. ISSN 0149-7634. doi: 10.1016/j.neubiorev.2013.02.002.
- Sharoni Shafir, Tom A. Waite, and Brian H. Smith. Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, 51 (2):180–187, January 2002. ISSN 1432-0762. doi: 10.1007/s00265-001-0420-8.
- William F. Sharpe. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*. *The Journal of Finance*, 19(3):425–442, 1964. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1964.tb02865.x.
- Na-Young So and Veit Stuphorn. Supplementary Eye Field Encodes Option and Action Value for Saccades With Variable Reward. *Journal of Neurophysiology*, 104(5):2634–2653, November 2010. ISSN 0022-3077. doi: 10.1152/jn.00430.2010.
- Aireza Soltani and Alicia Izquierdo. Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, 20(10):635–644, October 2019. ISSN 1471-0048. doi: 10.1038/s41583-019-0180-y.

-
- Mehran Spitmaan, Emily Chu, and Alireza Soltani. Salience-Driven Value Construction for Adaptive Choice under Risk. *The Journal of Neuroscience*, 39(26):5195–5209, June 2019. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2522-18.2019.
- William R. Stauffer, Armin Lak, and Wolfram Schultz. Dopamine Reward Prediction Error Responses Reflect Marginal Utility. *Current Biology*, 24(21):2491–2500, November 2014. ISSN 0960-9822. doi: 10.1016/j.cub.2014.08.064.
- Colin M. Stopper and Stan B. Floresco. Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, & Behavioral Neuroscience*, 11(1):97–112, March 2011. ISSN 1531-135X. doi: 10.3758/s13415-010-0015-9.
- Jonathan A. Sugam, Michael P. Saddoris, and Regina M. Carelli. Nucleus Accumbens Neurons Track Behavioral Preferences and Reward Outcomes During Risky Decision Making. *Biological Psychiatry*, 75(10): 807–816, May 2014. ISSN 0006-3223. doi: 10.1016/j.biopsych.2013.09.010.
- Jung Hoon Sul, Suhyun Jo, Daeyeol Lee, and Min Whan Jung. Role of rodent secondary motor cortex in value-based action selection. *Nature Neuroscience*, 14(9):1202–1208, September 2011. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.2881.
- Philippe N. Tobler and Elke U. Weber. Valuation for Risky and Uncertain Choices. *Neuroeconomics: Decision Making and the Brain: Second Edition*, pages 149–172, September 2013. doi: 10.1016/B978-0-12-416008-8.00009-7.
- Agnieszka Tymula, Lior A Rosenberg Belmaker, Lital Ruderman, Paul W Glimcher, and Ifat Levy. Like cognitive function, decision making across the life span shows profound age-related changes. *Proceedings of the National Academy of Sciences*, 110(42):17143–17148, 2013.
- Anne E. Urai, Anke Braun, and Tobias H. Donner. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications*, 8(1):14637, March 2017. ISSN 2041-1723. doi: 10.1038/ncomms14637.
- Mieke van Holstein and Stan B. Floresco. Dissociable roles for the ventral and dorsal medial prefrontal cortex in cue-guided risk/reward decision making. *Neuropsychopharmacology*, 45(4):683–693, March 2020. ISSN 0893-133X, 1740-634X. doi: 10.1038/s41386-019-0557-7.
- Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, September 2017. ISSN 0960-3174, 1573-1375. doi: 10.1007/s11222-016-9696-4.
- Vinod Venkatraman, John W. Payne, and Scott A. Huettel. An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence. *Organizational Behavior and Human Decision Processes*, 125(2):73–87, November 2014. ISSN 0749-5978. doi: 10.1016/j.obhdp.2014.06.003.
- John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University press, Princeton (N.J.), 3rd ed edition, 1953. ISBN 978-0-691-04183-4.
- Bethany J. Weber and Gretchen B. Chapman. Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes*, 97(1):31–46, May 2005. ISSN 0749-5978. doi: 10.1016/j.obhdp.2005.03.001.
- Allan F Williams. Teenage drivers: Patterns of risk. *Journal of Safety Research*, 34(1):5–15, January 2003. ISSN 00224375. doi: 10.1016/S0022-4375(02)00075-0.
- S.-W. Wu, M. R. Delgado, and L. T. Maloney. The Neural Correlates of Subjective Utility of Monetary Outcome and Probability Weight in Economic and in Motor Decision under Risk. *Journal of Neuroscience*, 31(24):8822–8831, June 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0540-11.2011.

Yan-Hua Xuan, Shu Li, Rui Tao, Jie Chen, Li-Lin Rao, X. T. Wang, and Rui Zheng. Genetic and Environmental Influences on Gambling: A Meta-Analysis of Twin Studies. *Frontiers in Psychology*, 8:2121, December 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.02121.

J. Frank Yates, editor. *Risk-Taking Behavior*. Wiley Series, Human Performance and Cognition. Wiley, Chichester, West Sussex, England ; New York, 1992. ISBN 978-0-471-92250-6.

Kelly A. Zalocusky, Charu Ramakrishnan, Talia N. Lerner, Thomas J. Davidson, Brian Knutson, and Karl Deisseroth. Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature*, 531(7596):642–646, March 2016. ISSN 0028-0836. doi: 10.1038/nature17400.