

# **Analyses of decisions under risk in rats**

by

**Xiaoyue Zhu**

A dissertation submitted in partial fulfillment of

the requirements for the degree of

**Doctor of Philosophy**

**Center for Neural Science**

**New York University**

**January 2022**

---

Jeffrey C. Erlich

© Xiaoyue Zhu

All Rights Reserved, 2021



## Dedication

*To my family and loved ones, for their unconditional faith and support*

## Acknowledgements

I am deeply indebted to my advisor Jeffrey C. Erlich. Without his guidance, patient mentoring, and unparalleled knowledge, none of the projects would have been conceived, let alone completed. The invaluable lessons I learned, from modeling techniques to a critical, quantitative way of thinking, will stay with me for my lifetime. I am lucky to be a part of his small yet creative team, which felt like an academic family that will always be supportive.

I wish to thank members of my committee – Paul Glimcher, Robert Froemke and Xinying Cai for their guidance and truthful feedback, especially on how to move forward when the original project became impractical to be completed. It was through this experience that I learned the value of letting go, and found courage to start a new project which turned out to be fruitful.

I am extremely grateful to all the technicians – Anyu Fang, Nengneng Gao and especially Yingkun Li, who trained hundreds of animals everyday and were always so diligent and dependable. I would like to thank Joshua Möller-Mara, who helped build the high-throughput system and developed the risky choice task. His knowledge of statistics and obsession with methodology bettered my analyses in so many ways. I wish to thank Sylvain Dubroqua, who helped shape rodent behavior in the risky choice task and taught me several surgical skills. Most importantly, his cheerful presence and *l'humour français* breathed life into my graduate years like the plants he so lovingly cared for. I am indebted to Evgeniya Lukinova, her human work contributed tremendously to CHAPTER 2. Despite being a mom and teaching professor, she always provided timely and meticulous feedback on my work – with such a strong work ethic I wish her all the best in her future endeavours. Special thanks to Chaofei

Bao, who helped collect inactivation results in CHAPTER 3. His academic humility and curiosity deeply inspires me. I am thankful for all the other graduate students in the lab – Liujunli Li and Jingjie Li, who fostered an open and stimulating intellectual environment, in which we all grew to be a better version of ourselves. I gratefully acknowledge the work of veterinaries in the facility, which ensured a humane and ethical environment for the animals. Lastly, as unconventional as it may be, I express my sincerest gratitude to all the animals sacrificed throughout the years. Creation of scientific knowledge bears real costs. I will be forever reminded of the lives sacrificed, no matter how small, to make this dissertation happen.

## Abstract

All interesting choices in life involve some degree of uncertainty. One form of uncertainty is risk: where the possible odds and outcomes are known. For example, when playing roulette in a casino, the probability of doubling your money on red is approximately 47.4%. But we can also learn from history: the chance of getting lung cancer is 20 times greater for regular smokers. Bet on red? Enjoy the smooth taste of a Lucky Strike? These are decisions under risk.

Decisions under risk have been extensively studies by economics, psychologists and neuroscientists. Still, many questions about the behavioral and neurological mechanisms underlying risky choice remain. Recently, advances in techniques have led to an increased focus on mice and rats as models for decision-making. Despite a rising adoption of rodents, there has been a gap between human-primate and rodent work on risky choice. While human and monkey research has explored both decisions under unexpected uncertainty (where environmental statistics may change suddenly) and risk (where the odds and outcomes are known), rodent work has largely focused on unexpected uncertainty. Using an automatic high-throughput training system, we developed two rodent tasks for investigations of choices under risk.

In the ‘risky choice’ task, mice and rats chose between a lottery and a small but guaranteed surebet. The lottery magnitude varied independently across trials, and was signalled by an auditory cue. Human data was also collected on a similar nonverbal task. Using a model-comparison approach, 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 there was considerable overlap across species in the

degree of risk-tolerance, animal subjects, compared to humans, were better described by models that included stimulus-independent biases and history dependence.

In rats trained on the risky choice task, we pharmacologically silenced two cortical regions: the frontal orienting field (FOF) and posterior parietal cortex (PPC). Although both regions were implicated in decision-making, inactivations produced strikingly different results. While FOF inactivations substantially biased the animals to choose the surebet, PPC inactivations produced minimal effects. With model-based analysis, we found that changes following FOF inactivation was best characterized by a decrease in  $\rho$ , the exponent of the utility function. Moreover, we constructed a biophysical model with FOF as part of a network for sensory-to-value transformation that parsimoniously explained our inactivation results.

In the ‘perceptual gambling’ task, rats made choices informed by both perceptual and value cues on a trial-by-trial basis. As in typical perceptual tasks, subjects were rewarded for correctly categorizing a tone relative to a learned threshold. To add an economic component, a light indicated whether correct responses to one side gave higher rewards than correct responses to the other side. On trials with some perceptual uncertainty, it could be worthwhile to shift choices to the unlikely option, if it had higher expected value. Using a mixture Bayesian decision theory model, we found that subjects tend to under-shift, which could be interpreted as being over-confident in their perceptual beliefs or as being risk-averse. We present the behavior as a promising template, for future researcher, to study percept-value integration and confidence during choices.

# Table of Contents

Dedication . . . . .	iv
Acknowledgements . . . . .	v
Abstract . . . . .	vii
List of Figures . . . . .	xvi
List of Tables . . . . .	xvii
<b>1 Introduction</b>	<b>1</b>
1.1 Types of decision making . . . . .	2
1.2 Models of risky choice . . . . .	5
1.3 Human vs. rodent tasks of risky choice . . . . .	10
1.4 Role of FOF and PPC in decision making . . . . .	13
1.5 Dissertation outline . . . . .	16
<b>2 An analysis of decision under risk in mice, rats and humans</b>	<b>19</b>
2.1 Introduction . . . . .	19
2.2 Materials and methods . . . . .	25
2.2.1 Animal behavior . . . . .	25
2.2.2 Human behavior . . . . .	29

2.2.3	Modeling risky choice . . . . .	31
2.2.4	Analysis . . . . .	36
2.3	Results . . . . .	41
2.3.1	A cross-species paradigm of risky choice . . . . .	41
2.3.2	Systematic modeling of risky choice . . . . .	44
2.3.3	Model-based estimation of risk attitudes . . . . .	47
2.3.4	Animals are less utility-maximizing than humans . . . . .	50
2.3.5	Estimation of rho is model-dependent . . . . .	54
2.4	Discussion . . . . .	57
2.5	Supplementary figures . . . . .	62
<b>3</b>	<b>Frontal but not parietal cortex is required for economic decisions under risk</b>	<b>75</b>
3.1	Introduction . . . . .	75
3.2	Materials and methods . . . . .	79
3.2.1	Subjects . . . . .	79
3.2.2	Behavior . . . . .	80
3.2.3	Experimental manipulations . . . . .	83
3.2.4	Analysis . . . . .	84
3.2.5	Surebet learning . . . . .	96
3.3	Results . . . . .	96
3.3.1	Behavior . . . . .	96
3.3.2	Effects of silencing FOF and PPC on the risky choice task . . .	97
3.3.3	A three-agent mixture model of risky choice . . . . .	102
3.3.4	Bilateral FOF inactivation reduced the utility exponent $\rho$ . . .	105

3.3.5	Bilateral PPC inactivation did not impair learning . . . . .	109
3.4	Discussion . . . . .	113
3.4.1	Role of FOF . . . . .	115
3.4.2	Role of PPC . . . . .	117
3.4.3	Conclusion . . . . .	120
3.5	Supplementary figures . . . . .	121
<b>4</b>	<b>A rodent paradigm for studying perceptual decisions under asymmetric reward</b>	<b>132</b>
4.1	Introduction . . . . .	132
4.2	Materials and methods . . . . .	135
4.2.1	Subjects . . . . .	135
4.2.2	Behavior . . . . .	135
4.2.3	Modeling with Bayesian decision theory . . . . .	140
4.2.4	Analysis . . . . .	147
4.3	Results . . . . .	153
4.3.1	The perceptual gambling task . . . . .	153
4.3.2	Behavior . . . . .	155
4.3.3	A three-agent mixture model with Bayesian decision theory . .	157
4.3.4	The mixture-BDT model is insufficient to account for subjects' behavior . . . . .	160
4.3.5	The animals are not reward-maximizing . . . . .	163
4.4	Discussion . . . . .	165
4.5	Supplementary figures . . . . .	172

<b>5 Conclusion</b>	<b>177</b>
5.1 General summary . . . . .	177
5.2 Limitations and recommendations . . . . .	179

# List of Figures

2.1	The risky choice task and animal behavior . . . . .	43
2.2	The human risk nonverbal task and behavior . . . . .	45
2.3	Systematic modeling of risky choice . . . . .	48
2.4	Fits by the rho-sigma model . . . . .	51
2.5	Fits by the history-mix-rho-kappa-sigma model . . . . .	52
2.6	Model comparison with 10-fold cross-validation . . . . .	55
2.7	Species correlation of $\rho$ estimates from all the models containing $\rho$ . .	57
2.8	Example behavioral performance over training . . . . .	62
2.9	The models can recover the data-generating parameters well . . . .	63
2.10	Example mouse's choices superimposed with predictions from the rho-sigma model . . . . .	64
2.11	Example rat's choices superimposed with predictions from the rho-sigma model . . . . .	65
2.12	Example binned human's choices superimposed with predictions from the rho-sigma model . . . . .	66
2.13	Example mouse's choices superimposed with predictions from the history-mix-rho-kappa-sigma model . . . . .	67

2.14 Example rat's choices superimposed with predictions from the history-mix-rho-kappa-sigma model . . . . .	68
2.15 Example binned human's choices superimposed with predictions from the history-mix-rho-kappa-sigma model . . . . .	69
2.16 MAP parameter estimation by the history-mixture-rho-kappa-sigma model for each mouse . . . . .	70
2.17 MAP parameter estimation by the history-mixture-rho-kappa-sigma model for each rat . . . . .	70
2.18 MAP parameter estimation by the history-mixture-rho-kappa-sigma model for each human subject . . . . .	71
2.19 Ten-fold cross-validation results comparing all 18 models for the mouse population . . . . .	72
2.20 Ten-fold cross-validation results comparing all 18 models for the rat population . . . . .	73
2.21 Ten-fold cross-validation results comparing all 18 models for the human population . . . . .	74
 3.1 The risky choice task and animal behavior . . . . .	98
3.2 Bilateral and unilateral infusions in FOF and PPC. . . . .	100
3.3 Unilateral PPC inactivation induced an ipsilateral bias on free trials .	103
3.4 The three-agent mixture model and model fits . . . . .	106
3.5 Bilateral FOF inactivation reduced the utility exponent . . . . .	110
3.6 Biophysical model of FOF silencing . . . . .	111
3.7 Bilateral PPC inactivations did not impair learning of new surebet magnitudes . . . . .	114

3.8	Timeline the experimental treatments for each animal . . . . .	121
3.9	Supplementary details of task and behavior . . . . .	122
3.10	Histology summary . . . . .	123
3.11	Effects of FOF inactivations for each animal . . . . .	124
3.12	Effects of PPC inactivations for each animal . . . . .	124
3.13	Reaction time analysis following bilateral FOF and PPC inactivations	125
3.14	Reaction time analysis following unilateral FOF and PPC inactivations	126
3.15	Supplementary details for the three-agent model . . . . .	127
3.16	Three-agent model fits on control sessions . . . . .	128
3.17	Inactivation model predictions . . . . .	129
3.18	Inactivation model fits on control and FOF inactivation sessions . . .	130
3.19	Behavioral adaptation in the surebet learning experiment . . . . .	131
4.1	The generative model of Bayesian inference . . . . .	141
4.2	The perceptual gambling task . . . . .	156
4.3	Animal behavior in the perceptual gambling task . . . . .	158
4.4	The mixture-BDT model . . . . .	161
4.5	The mixture-BDT model fits . . . . .	163
4.6	The four-parameter sigmoid model fits . . . . .	164
4.7	The animals were sub-optimal in obtaining reward . . . . .	166
4.8	Supplementary details for the mixture-BDT model . . . . .	172
4.9	Animal's choices were influenced by the outcome from the previous trial	173
4.10	The mixture-BDT model fits for each subject . . . . .	174
4.11	The mixture-BDT model only allows for symmetric and horizontal shifts as a result of asymmetric reward . . . . .	175

4.12 The four-parameter sigmoid model fits for each subject . . . . . 176

# List of Tables

3.1	Fits from the three-agent inactivation model . . . . .	107
4.1	The action cost table . . . . .	144
4.2	The perceptual difficulty and reward multiplier effective for each animal	155

# Chapter 1

## Introduction

“Excellence is never an accident. It is always the result of high intention, sincere effort, and intelligent execution; it represents the wise choice of many alternatives. Choice, not chance, determines your destiny.”

— Aristotle

The act of choice is perhaps the most ordinary yet consequential aspect of human life. From choosing an ice cream flavor to a career path, the power to choose is considered as a hallmark of freewill by many philosophers (O’Connor and Franklin, 2021). In a liberalized society, a choice not only affects one’s material and psychological welfare, but can shape policy-making and financial markets in all their forms, affecting other individuals, groups, and institutions that reside within a highly interconnected ecosystem (Moisander et al., 2010). When society at large is fueled by an intertwined web of deliberate and influenced choices, understanding the process of how a choice is made, known as decision making, has never appeared more urgent.

From a public health perspective, studying mechanisms of choice is of substantial

interest. Maladaptive decision making has been observed in people with anxiety (Maner et al., 2007), schizophrenia (Heerey et al., 2008), substance abuse disorder (Hoffman et al., 2008), depression (Harlé et al., 2010), and many other psychiatric disorders (see Lee, 2013, for a review). While some evidence suggest that genetics partially accounts for variations in individual choices (23and Me Research Team et al., 2019, Xuan et al., 2017, Kawamura et al., 2013), external factors such as poverty (Carvalho et al., 2016, Bartoš et al., 2021) and stress (Kandasamy et al., 2014, Koot et al., 2013) can also have an impact. All of this indicates that decision making is a complex process that employs multiple brain circuits and is state- and context-dependent. As such, animal models that allow for systematic investigation of robust decision-making behavior are key to unravelling the connection between genes, neural circuits and decisions.

This chapter provides a general background on tasks and models developed by decision science, with a special focus on choice under risk. To motivate my research topics, I will briefly review the status of rodent literature in relation to human literature, and discuss the importance of two cortical regions in decision making.

## 1.1 Types of decision making

Decision making is an abstract term referring to the integration and transformation of external information with internal states into an action. It encompasses an incredibly wide range of human activities, from judging the color of traffic lights and deciding where to eat lunch, to complex sociopolitical policy-making. The field in neuroscience that studies decision-making behavior is traditionally composed of

two groups. The first, perceptual decision making, is concerned with how animals and humans detect, discriminate, and categorize information from noisy sensory inputs (Hanks and Summerfield, 2017). The field began as an intellectual child of psychophysics, which studied people’s judgments of a simple sensory features (e.g. brightness, loudness). The key to perceptual decisions is to convert noisy sensory information into a categorical, usually binary motor response. For example, a classic paradigm for studying such decisions in macaque monkeys is the random dot task (RDT) developed by Newsome and Pare (1988). Presented with a screen full of moving dots, subjects were instructed to report the general moving direction by making a saccade to a spatial target. The difficulty of the task was controlled by motion coherence: the percentage of dots moving in the same direction. Thus, by altering motion coherence, the researchers were able to study how monkeys integrate perceptual evidence over time, which also provided a robust paradigm for studying the neural substrates underlying perceptual decisions.

Rodents offer several advantages over the use of non-human primates for the study of decision-making. Apart from reduced ethical concerns and increased cost-effectiveness, there is a growing availability of molecular and genetic tools for recording and manipulating neural activity with high temporal and spatial precision (e.g. Deisseroth, 2014, Kramer et al., 2013). Moreover, rodents allow for semi-automated high-throughput training, allowing data to be collected efficiently (Brunton et al., 2013). A key rodent task for studying perceptual decisions is the Poisson Clicks Task, in which the rats reported the side which played more auditory clicks by poking into the left or right port (Brunton et al., 2013). The discrete pulses generated from a Poisson process allowed experimenters to quantitatively link cortical neural

activity to the number of pulses, demonstrating that the rodent brain is also capable of evidence accumulation as primates (Hanks et al., 2015).

On the other hand, the second group studies economic decisions, defined as the multi-attribute decisions that reveal an individuals' preferences (Padoa-Schioppa, 2011). This contrasts with perceptual decisions where there are *objectively* correct answers, e.g. choosing the side with more auditory clicks. In an economic choice, there is only a *subjectively* preferred answer which is revealed by choice. Does Chicago or New York have better pizza? While loyalists in each camp might claim there is a correct answer to that question, as impassive scientists we know it is a matter of preference. Here, I briefly describe three forms of economic choice: choice under certainty, choice under risk, and choice under uncertainty.

Choice under certainty involves the subject choosing from a set of options, each with a single deterministic outcome. This is studied in the laboratory by asking subjects to choose from a set of goods with different intrinsic values, such as: rats choosing between different quantities of banana- and bacon-flavored pellets (Gardner et al., 2017, 2018, 2019); monkeys choosing between varied amounts of fruit juice and peppermint tea (Padoa-Schioppa and Assad, 2006, 2008, Cai and Padoa-Schioppa, 2012); and humans choosing between candy bars and salty snacks (Chib et al., 2009); or between money, snacks and electronics (FitzGerald et al., 2009).

Choice under risk, on the other hand, involves choosing from a set of options, each with a well-defined probability for each outcome. Choosing between option A (50% chance of winning \$100 and 50% of winning \$0) and option B (100% chance of winning \$50) is a typical example. This is the main form of economic choice examined in this dissertation, and it will be described with details in SECTION 1.2 and SECTION 1.3.

Next, choice under uncertainty refers to choices made without explicit knowledge of the possible outcome, either of the outcome itself or the probability associated with it, with the latter referred as ambiguity (Knight, 1921). In fact, this is the form of choice that closely resembles some of the most important choices we make in real life. From deciding on a college major to making a financial investment, we do not know the exact outcomes and probabilities at the time of the decision, yet we are perfectly capable of making it under incomplete information. Interestingly, humans have a tendency to avoid ambiguity and prefer risky options where the probabilities are given, a phenomenon known as ambiguity-aversion (Heath and Tversky, 1991, Lauriola and Levin, 2001). Perhaps unsurprisingly, neuroimaging studies have revealed that different brain regions are engaged when making decisions under risk and under ambiguity, suggesting that they call for distinct cognitive mechanisms (Huettel et al., 2006).

Lastly, I wish to emphasize that this is by no means a complete list of all decision-making behavior. Other types such as intertemporal choice, effort-based choice, decisions in social contexts, and strategic choice are also ubiquitous in real-life and have proved to be invaluable research topics.

## 1.2 Models of risky choice

Choice under risk is one of the most widely researched forms of choice. As the experimenter can systematically vary the payout and probability of the gamble, it provides fertile ground for quantitative analysis using normative and descriptive models. The dominant normative model is the expected utility theory (EUT) formalized by Von Neumann and Morgenstern (1953), which assumes that a rational agent *should*

maximize her utility after considering all payouts and probabilities associated with a gamble. Utility, a pivotal concept in economics, refers to the subjective satisfaction after receiving or consuming a good or service. The expected utility (EU) can be computed as the average of the utilities of different outcomes weighted by their probabilities. Given a gamble  $X$ , we have

$$EU(X) = \sum_x p(x)u(x), \quad (1.1)$$

$$u(x) = f(v(x)) \quad (1.2)$$

where  $v(x)$  is the payoff of each outcome,  $p(x)$  is the respective probability,  $u(x)$  refers to the subjective utility after receiving payoff  $v(x)$ , transformed with a utility function  $f(v(x))$ . Although this function can take on many forms in theory, an exponential function  $u(x) = v(x)^\rho$  became the standard such that the exponent  $\rho$  alone determines the shape of it. In this framework, the decision maker's attitude towards risk, known as risk attitude, is synonymous with  $\rho$ . When  $\rho = 1$ , the decision maker is risk-neutral, as she would choose indifferently between receiving a guaranteed  $x$  (surebet) and receiving  $2x$  or nothing with equal probabilities. When  $\rho < 1$ , the utility function becomes concave (decelerating) and the decision maker would prefer the guaranteed  $x$ , as the expected utility of the lottery is now below the surebet, showing risk-aversion. Finally, when  $\rho > 1$ , the utility function becomes convex (accelerating) and the decision maker is risk-seeking. Importantly, an individual is considered *rational* if she makes decisions strictly following the theorem of utility maximization, regardless the shape of her utility function.

Observing that the curvature of the utility function alone can describe individual risk attitude, Arrow (1965) and Pratt (1964) derived a general definition of risk aversion that is independent of the utility's functional form:

$$ARA_u(x) = -u''(x)/u'(x) \quad (1.3)$$

where  $u'$  and  $u''$  represent the first and second derivative of the utility function, respectively. This measure of risk aversion, known as the constant absolute risk aversion (CARA), is complemented by other variants such as the constant relative risk aversion (CRRA), which recognizes that one's risk attitude may change based on her wealth level.

In the field of finance, the mean-variance model proposed by Markowitz (1968) treated risk as an inherent property of outcome distribution rather than a feature of the utility function. To understand the conflicting objective of high profit and low risk in a portfolio, Markowitz formulated the following equation:

$$WTP(X) = V(X) - \kappa R(X) \quad (1.4)$$

It specifies that people's willingness-to-pay (WTP) for a certain gamble  $X$  is a trade-off between its expected value  $V(X)$  and its variance  $R(X)$ , mediated by the parameter  $\kappa$  that captures the individual risk attitude, or more precisely, variance-aversion. A positive  $\kappa$  denotes risk-aversion whereas a negative  $\kappa$  denotes risk seeking. It must be noted that Markowitz himself is a disciple of EUT, as he believes that a rational decision *should* be characterized by the maximization of utility. This was precisely

the intention behind the mean-variance theory: it was meant to serve as a practical approximation of the individual investors' unknown utility functions, especially when the decision is to be made by one institutional investor (Markowitz, 2014). In fact, Levy and Markowitz (1979) have shown that a wide class of concave utility functions can be approximated using the mean-variance model with Taylor expansion. Despite their influence and normative strengths, experimental evidence had cast doubt on EUT and the mean-variance theory, showing that human choice behavior often violated axioms of the theory, with the most famous being the Ellsberg paradox (Ellsberg, 1961) and Allais paradox (Allais, 1953).

Prospect theory by two eminent psychologists Kahneman and Tversky (1979), on the other hand, was developed as a descriptive model to *explain* the observed choice behavior and violations of EUT. The key insight was that humans choose in relation to an internal reference point, which was generally assumed to be the current level of wealth, although its nature and formation is an ongoing research area (van Osch et al., 2006, Hunter and Gershman, 2018). It was postulated that humans choose with a concave utility function in the domain of gains, and a convex utility function in the domain of losses. In other words, they are risk-averse to gains and risk-seeking for losses. Additionally, prospect theory posits that people tend to underweight high probabilities and overweight low probabilities, a phenomenon known as probability weighting.

Since its inception in 1979, this seminal theory has been shown to account for many aspects of human choice. For example, the endowment effect – that the participants would ask for a higher price to sell an item than the price they previously acquired it for, can be attributed to loss aversion (Kahneman et al., 1990). In finance,

investors sometimes overprice positively skewed stocks despite negative average returns. Probability weighting offers an explanation for this behavior, that investors may overweight the low probability of a certain stock becoming the “next Google” (Barberis and Huang, 2008). Prospect theory again found its voice in explaining the disposition effect, that traders tend to sell stocks generating a profit and hold those making a loss (Frazzini, 2006). This puzzling behavior can be attributed to the convex utility function in the loss domain, as the trader holds the stock in hope of recovering his loss in the future. Recently, it has been tested in rats choosing under different lottery payouts and probabilities. It was found that the rats exhibit risk-aversion and probability weighting, consistent with predictions from the theory (Constantinople et al., 2019b).

As described above, there are a myriad of models from economics, finance, psychology, and ecology (see Kacelnik and Bateson, 1996), each examining risk through a lens assembled with assumptions. This poses a considerable challenge to a neuroscientist who wishes to find neural correlates of risk attitude, as it may be the exponent on the utility function, the degree of variance aversion, or even the degree to which a heuristic strategy is adopted (Brandstätter et al., 2006). It is further complicated by the fact that choices are often stochastic, i.e. the decision maker may respond differently given the same options each time. The stochasticity may come from a noisy representation of one’s utility, or it may be due to noise after the comparison has been made (Bhatia and Loomes, 2017). Undoubtedly, the measured risk attitude will not reflect the true risk attitude if the potential source of noise is not considered appropriately. To better account for choice stochasticity, the scalar utility theory (SUT) proposed by Kacelnik and Brito e Abreu (1998) assumes a utility distribution

that follows Weber's law – the noise scales with magnitude of the mean. SUT has had success in explaining decision-making behavior in starlings (Marsh, 2002) and honeybees (Shafir et al., 2002).

### 1.3 Human vs. rodent tasks of risky choice

Generally speaking, there are two types of tasks a researcher may use to study risky choice in humans. The first type involves presenting full information, usually described in text, so that the participant has full knowledge of the outcomes and probabilities associated with each option. A typical example would be:

Choose between

Option A: \$250,000 with certainty

Option B: 50% chance of \$500,000 and 50% chance of \$0

This is known as decision *from description*, and has been widely adopted in economics and psychology experiments (e.g. Camerer, 1989). However, decisions in real life rarely have well-described outcomes. It is often the case that people rely on their previous experience, specifically learned probabilities and outcomes, to make decisions (Hertwig et al., 2004). Researchers interested in studying risky choice *from experience* resort to one of the following experimental tools: the sampling paradigm, the full-feedback paradigm, and the partial-feedback paradigm. As the name suggests, participants learn the outcome contingencies either by sampling prior to the 'real deal' (Weber et al., 2004), by receiving feedback on every trial (Yechiam and Busemeyer, 2006), or receiving the final payoff only (Erev and Barron, 2005). Interestingly, a *description-experience gap* has emerged. People tend to underweight

outcomes with low probabilities when making decisions from experience, opposite to prospect theory’s prediction. When making decisions from description, however, people overweight outcome with high probabilities, consistent with prospect theory (Hertwig and Erev, 2009). Although an elaboration of the description-experience gap is beyond the scope of this section, it is important to be aware of it as making decisions from description and from experience may employ distinct cognitive mechanisms (Hertwig et al., 2018). In particular, decisions from experience may engage *remembered utility*, a memory signal representing the outcome from previous occasions (Berridge and O’Doherty, 2014), whereas decisions from description may engage *predicted utility*, which was first defined by Kahneman et al. (1999) as “beliefs about the experienced utility of outcomes”.

Regardless of making decisions from description or experience, one characteristic the human tasks share is that participants choose under ‘expected uncertainty’ – an idea that there is known unpredictability in the environment (Dayan and Yu, 2003). In contrast, ‘unexpected uncertainty’ refers to the idea that the statistics underlying the environment may suddenly change, thus requiring updating of previously held beliefs (Dayan and Yu, 2003, Nassar et al., 2010). Converging evidence suggests that expected and unexpected uncertainty may recruit distinct brain circuits mediated by acetylcholine and norepinephrine, respectively (Dayan and Yu, 2003, Payzan-LeNestour et al., 2013, Soltani and Izquierdo, 2019). Having established the distinction, it will become clear why the majority of current rodent tasks are more befitting to the study of unexpected uncertainty than risky choice.

The workhorse of such rodent tasks is the probability discounting task (e.g. St Onge and Floresco, 2009, Larkin et al., 2016), in which food-deprived rats choose

between a small and certain (1 pellet) and a large but risky (0 or 4 pellets) option by pressing a lever. The probability of reward from the risky option systematically changes in a block-based manner, for example, from 100% to 12.5% either in a descending or ascending order. Importantly, the sudden change in reward probability is not explicitly cued – the animal has to infer the probability by keeping track of their choices and reward obtained instead. Using the probability discounting task, Floresco and colleagues identified regions closely associated with the learning and updating of action-values, in particular the amygdala (Larkin et al., 2016, Jenni et al., 2017, Bercovici et al., 2018), basal ganglia (Stopper et al., 2013, Floresco et al., 2018), and orbital and medial prefrontal areas (Onge et al., 2011, van Holstein and Floresco, 2020).

Another similar rodent task extends the outcome of the risky option to the domain of punishment. Coined as the risky decision-making task, animals choose between an option that delivers a small and guaranteed reward, and an option that delivers a large reward which is accompanied by a certain probability of footshock (Simon et al., 2009, Cooper et al., 2014, Shimp et al., 2015). The probability of the footshock either increases or decreases in consecutive blocks as in the probability discounting task, without any explicit cues. It can be readily seen that these rodent tasks are in fact eliciting behavior under unexpected uncertainty rather than expected uncertainty, as the rats have to make decisions under incomplete information of the outcomes. Consequently, these rodent results are not directly comparable to those from humans, where risky choice *per se* was the subject of investigation.

The only known studies of expected uncertainty to date is by Constantinople et al. (2019b,a), where rats made well-informed decisions under explicit audiovisual

cues on a trial-by-trial basis. On each trial, auditory clicks were played from the left and right speaker where its click rate indicated the payout magnitude. At the same time, the number of light flashes from the left and right choice port signalled the payout probability. By introducing a range of lottery magnitudes and probabilities, the experimenters were able to construct rats' psychometric curves and found that not only were they risk-averse, they also exhibited probability distortion as predicted by prospect theory. Demonstrating that rats are capable of making utility-maximizing decisions guided by complex cues, Constantinople et al.'s task announced the arrival of a new chapter of studying risky choice in rodents.

## 1.4 Role of FOF and PPC in decision making

Two inherent properties of risky choice are outcome value and outcome uncertainty. Consequently, studies looking for neural representation of risky choice will inevitably find correlates of either outcome value, uncertainty, or a combination of the two. As value and uncertainty are ubiquitous signals in the brain, detected in striatum (Schultz et al., 2008, Yacubian et al., 2007, Kable and Glimcher, 2007), insula (Burke and Tobler, 2011, Preuschoff et al., 2008), orbitofrontal cortex (Padoa-Schioppa and Assad, 2006, Van Duuren et al., 2009), lateral prefrontal cortex (Tobler et al., 2009, Kennerley et al., 2009), anterior cingulate cortex (Hayden and Platt, 2010, Levy et al., 2010), and even hippocampus (Harrison et al., 2006), a full review of the neural representation of risky choice is beyond the scope of this section (see Bach and Dolan, 2012, Platt and Huettel, 2008, Orsini et al., 2015, for excellent reviews). Instead, I will exclusively focus on two regions here: the frontal orienting field

(FOF) and the posterior parietal cortex (PPC), two rodent regions as well as their primate homologs have been implicated in both perceptual and economic choices.

The FOF is a putative rodent homolog of primate frontal eye field (FEF, Erlich et al., 2011, Hanks et al., 2015), also known by M2, medial agranular cortex and medial precentral cortex (Barthas and Kwan, 2017). It receives connections from the visual, auditory, somatosensory, parietal, and orbital areas, while reciprocally projecting to many prefrontal and parietal cortices (Condé et al., 1995, Corwin and Reep, 1998). Due to its anatomical connectivity, the FOF has been described as ‘a key link between multimodal sensory inputs and organized motor output’ (Reep et al., 1987). Experiments have found that FOF activity not only correlates with reward value (Sul et al., 2011), but also flexibly adapts based on stimulus-action contingencies (Siniscalchi et al., 2016), suggesting that the ability to use learned sensorimotor associations to guide actions could be a central feature of the FOF. Using a memory-guided frequency discrimination task, Erlich et al. (2011) recorded from the FOF and found early choice-related activity with trial-to-trial variations matching the upcoming response, establishing for the first time the involvement of FOF in decision making. In a follow-up experiment using the Poisson clicks task, Erlich et al. (2015) showed that unilateral silencing of the FOF biased the animals to choose the side with infusion, and this bias is larger for trials with longer memory periods (Piet et al., 2017). Bilateral silencing also produced an impairment whose intensity was dependent on the delay of integration. Overall, these results support the hypothesis that FOF serves as a bottleneck for higher order cognitive processes to guide decisions in rats.

Furthermore, the FOF’s primate homolog FEF, has been implicated in choices

under risk. Using an oculomotor gamble task where monkeys chose between pairs of gambles by making eye movements to different targets, Chen et al. (2020) showed that FEF activity was modulated by the expected reward of different eye movements before, during and even after the targets appeared. The same group also tested the causal role of supplementary eye field (SEF), an oculomotor area that shares many afferent and efferent projections with the FEF (Chen and Wise, 1995). They found that SEF inactivation significantly reduced the frequency of risky choices in monkeys, i.e. they chose the option with higher variance less, even when both options had the same expected value. Moreover, the reduction was best explained by a decreased risk preference term (Chen and Stuphorn, 2018). Thus, evidence from rodent and monkey experiments suggest that the FOF may be a critical locus for risky choice.

On the other hand, the PPC is an associative area located between the visual and somatosensory cortices. In both primates and rodents, it is reciprocally connected with striatum, visual, auditory, somatosensory, orbital, and frontal motor areas, making it ideally suited for the integration of sensorimotor information (Whitlock, 2017). Evidence from primate and rodent studies heavily implicated PPC in perceptual decisions. Using the random dot task, early researchers recorded from a subregion in monkey PPC and discovered that its activity correlated with evidence strength, predicting the monkey’s upcoming choice (Shadlen and Newsome, 2001, Roitman and Shadlen, 2002). In rat PPC, Hanks et al. (2015) observed such evidence-dependent activity in a memory-guided Poisson clicks task. Raposo et al. (2014) also found choice activity in rats performing a multisensory decision task. Moreover, choice-specific activity has been observed in mouse PPC in a visual discrimination task (Pho et al., 2018), and in a virtual navigation decision-making task (Harvey et al., 2012).

Surprisingly, given the consistent decision-related signatures found in PPC, causal manipulations of the region produced minimal impairment on perceptual choices in monkeys (Katz et al., 2016) and rats (Erlich et al., 2015). However, PPC inactivation reliably biased ‘free choice’ trials, where the animal was rewarded regardless of the choice, such that it simply chose its preferred side (Erlich et al., 2015, Katz et al., 2016). One key difference between the perceptual choice and the free choice is that the former is externally guided by sensory cues, whereas the latter is internally guided by the animal’s side bias. It is possible that the PPC is involved in generating actions pertaining to one’s internal preference only. Risky choice, as a type of economic choice, is characterized by the expression of individual risk preference. Furthering the argument, signatures of expected value (Platt and Glimcher, 1999) and subjective action value (Dorris and Glimcher, 2004) have been found in PPC. These results support a role of PPC in the dynamic unfolding of internally guided decisions, such as risky choice.

## 1.5 Dissertation outline

So far, I have attempted to provide a general picture of the extraordinarily diverse field that is decision science, with an emphasis on risky choice. With the rapid advancement of data processing and modeling techniques, I argue that the future of decision science must be highly interdisciplinary and data-centric – rather than be bound by field-specific assumptions, we should let data speak for itself. During my years in Erlich lab at NYU Shanghai, together as a team, we have developed an automated high-throughput animal training system. Each training apparatus features

an eight-port pokewall, two speakers, and a fully automated data pipeline, allowing us to train animals on complex behavioral tasks and collect millions of trials. With this powerful system, I was able to train animals on two decision-making behaviors involving risk: the risky choice task and the perceptual gambling task. The next three chapters will present results from these behavioral paradigms.

In Chapter 2, I will introduce the risky choice behavior collected from rats and mice in the high-throughput system, as well as human behavior on a similar non-verbal task (developed and collected by Evgeniya Lukinova). The risky choice task differs from traditional rodent tasks in that the subject makes well-informed decisions on a trial-by-trial basis, choosing under expected uncertainty. To understand which model captures animal and human data the best, I fit a collection of models inspired by the expected utility theory, the mean-variance model, the scalar utility theory, and the heuristic-based models. Results of the modeling suggest there is a fundamental difference between human and animals in terms of the strategy used to perform the task. This task and modeling framework provides a foundation for rigorous circuit level exploration of the neurobiology of risky choice.

In Chapter 3, I will present results from silencing the FOF and PPC with muscimol on the risky choice behavior. Interestingly, although neurons in both frontal and parietal cortex are known to encode decision variables, pharmacological inactivation produced strikingly different results. While FOF inactivations substantially biased the animals to choose the surebet, both unilateral and bilateral PPC inactivations produced minimal effects. Using model-based analysis, we found that changes following FOF inactivation was best characterized by a decrease in  $\rho$ , the exponent of the utility function. Moreover, we constructed a biophysical model with FOF as part of

the network and provided a mechanistic account for the value-to-utility transformation.

As described in SECTION 1.1, the neural mechanisms underlying perceptual decisions have typically been examined separately from those underlying economic decisions. In Chapter 4, I will present the perceptual gambling task, which requires the subjects to integrate perceptual and value information to make a choice. Animals learned the task after a parameter-sensitive training process. I then fit the behavior using a model developed with Bayesian decision theory and found that it was insufficient in capturing certain aspects of behavior. I will also discuss how future researchers can improve the behavior and the promising research topics it enables, including percept-value integration and the study of confidence.

# Chapter 2

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

### 2.1 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 ( $\sim 30\%$ ) 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.

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.

## **2.2 Materials and methods**

### **2.2.1 Animal behavior**

#### **2.2.1.1 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.

#### **2.2.1.2 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 2.1A), 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.

### 2.2.1.3 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  $1.96 \pm 0.85$ mL rewards in a session, whereas a rat obtained  $6.7 \pm 2.6$ mL.

#### **2.2.1.4 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  $129 \pm 56$  days to complete training and have their sessions meet the criteria of inclusion, whereas rats required only  $90 \pm 39$  days.

#### **2.2.1.5 Inclusion criteria**

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.

### 2.2.2 Human behavior

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

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

Then subjects had 20-22 training trials and continued with the real task, where each coin was worth money (the exact conversion rate differed from session to session since the distribution of the acquired coins for each session was rescaled to resemble a \$5-\$20 range for the bonus payment, on average subjects in the nonverbal task got bonus of \$12). There were three versions of the task during the initial session. Each session followed the same timeline (FIGURE 2.2A). First, subjects heard the sound cue (2 seconds) and saw two circles (blue for surebet and yellow for lottery option). Then, 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. At the same time, subjects saw a visual of money and money bag and heard rewarding coin dropping sound, if they got reward in this trial. The choice set from the first version is in FIGURE 2.2C, LEFT (22 training + 180 real trials). In the second version (compared to the first version), we added images of coins of what a subject would have earned, even if he lost the lottery on a given trial. We also added six forced trials in the beginning of the training phase. In the third version (compared to the first version), we changed the choice set to the one in FIGURE 2.2C, RIGHT and added six forced risk trials in the beginning of the training phase (20 training + 216 real trials). We followed up with 50 subjects who showed at least some sensitivity in their choices, tested by a generalized linear model (GLM):

$$\text{chose\_lottery} \sim \text{lott\_mag} + \text{lott\_prob} \quad (2.1)$$

where `chose_lottery` is 1 if lottery was chosen on a trial, `lott_mag` is the lottery

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

### 2.2.3 Modeling risky choice

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

#### 2.2.3.1 History-independent models

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.

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

$$U_L = V_L^\rho P_L \quad (2.2)$$

$$U_{SB} = V_{SB}^\rho \quad (2.3)$$

$$\hat{U}_L \sim \mathcal{N}(U_L, \sigma) \quad (2.4)$$

$$\hat{U}_{SB} \sim \mathcal{N}(U_{SB}, \sigma) \quad (2.5)$$

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (2.6)$$

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  $\sigma$ . 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:

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (2.7)$$

$$= 1 - \Phi(0; V_L^\rho P_L - V_{SB}^\rho, \sqrt{2}\sigma) \quad (2.8)$$

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.

**The kappa-sigma model** The kappa-sigma model models risk preference  $\kappa$  as the degree of variance aversion:

$$\hat{U}_L \sim \mathcal{N}(P_L(V_L - \kappa\sqrt{(V_L)}), \sigma) \quad (2.9)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}, \sigma) \quad (2.10)$$

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (2.11)$$

where  $(V_L)$  denotes the variance of the lottery option.

**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:  $\rho$  for the curvature of utility function, and  $\kappa$  for the degree of variance aversion:

$$\hat{U}_L \sim \mathcal{N}(P_L(V_L^\rho - \kappa\sqrt{(V_L^\rho)}), \sigma) \quad (2.12)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}^\rho, \sigma) \quad (2.13)$$

$$P(\text{Choose Lottery}) = P(\hat{U}_L > \hat{U}_{SB}) \quad (2.14)$$

**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  $\gamma$  in  $\mathcal{N}(EU, \gamma EU)$  was estimated instead of  $\sigma$  in the non-scalar models.

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

$$\hat{U}_L \sim \mathcal{N}(P_L V_L^\rho, \gamma P_L V_L^\rho) \quad (2.15)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}^\rho, \gamma V_{SB}^\rho) \quad (2.16)$$

For the kappa-scalar model:

$$\hat{U}_L \sim \mathcal{N}(P_L \left( V_L - \kappa \sqrt{(V_L)} \right), \gamma P_L \left( V_L - \kappa \sqrt{(V_L)} \right)) \quad (2.17)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}, \gamma V_{SB}) \quad (2.18)$$

For the rho-kappa-scalar model:

$$\hat{U}_L \sim \mathcal{N}(P_L \left( V_L^\rho - \kappa \sqrt{(V_L)} \right), \gamma P_L \left( V_L^\rho - \kappa \sqrt{(V_L)} \right)) \quad (2.19)$$

$$\hat{U}_{SB} \sim \mathcal{N}(V_{SB}^\rho, \gamma V_{SB}^\rho) \quad (2.20)$$

**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  $\vec{\omega}$ :

$$\vec{P} = [P_{rational}, 1, 0] \quad (2.21)$$

$$\vec{\omega} = [\omega_{rational}, \omega_{lottery}, \omega_{surebet}] \quad (2.22)$$

$$\sum \vec{\omega} = 1 \quad (2.23)$$

$$P(\text{Choose Lottery}) = \vec{P} \cdot \vec{\omega} \quad (2.24)$$

The history-independent three-agent mixture models include: mix-rho-sigma, mix-kappa-sigma, mix-rho-kappa-sigma, mix-rho-scalar, mix-kappa-scalar, and mix-rho-kappa-scalar. Note that they only differ in how the model specifies the rational agent.

### 2.2.3.2 History-dependent models

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,

$$\vec{P} = [P_{rational}, 1, 0] \quad (2.25)$$

$$\vec{\omega}_{win} = [\omega_{rational}^{win}, \omega_{lottery}^{win}, \omega_{surebet}^{win}] \quad (2.26)$$

$$\vec{\omega}_{lose} = [\omega_{rational}^{lose}, \omega_{lottery}^{lose}, \omega_{surebet}^{lose}] \quad (2.27)$$

$$\vec{\omega}_{SB} = [\omega_{rational}^{SB}, \omega_{lottery}^{SB}, \omega_{surebet}^{SB}] \quad (2.28)$$

$$\sum \vec{\omega}_{win} = \sum \vec{\omega}_{lose} = \sum \vec{\omega}_{SB} = 1 \quad (2.29)$$

$$P(\text{Choose Lottery}) = \begin{cases} \vec{P} \cdot \vec{\omega}_{win}, & \text{if lottery-win} \\ \vec{P} \cdot \vec{\omega}_{lose}, & \text{if lottery-lose} \\ \vec{P} \cdot \vec{\omega}_{SB}, & \text{if surebet} \end{cases} \quad (2.30)$$

The history-dependent three-agent models include: history-mix-rho-sigma, history-mix-kappa-sigma, history-mix-rho-kappa-sigma, history-mix-rho-scalar, history-mix-kappa-scalar, and history-mix-rho-kappa-scalar.

#### 2.2.4 Analysis

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).

#### 2.2.4.1 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:

$$\text{efficacy} = 0.5 \frac{\text{animal} - \text{random}}{\text{maximizer} - \text{random}} + 0.5 \quad (2.31)$$

#### 2.2.4.2 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.

#### 2.2.4.3 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 psycho-

metric curves as in FIGURE 2.1E-F and FIGURE 2.2D, we specified a GLM where the probability of a lottery choice was a logistic function of  $EV_{lottery} - EV_{surebet}$ . The expected value of lottery is the product of the lottery magnitude and lottery probability ( $EV_{lottery} = P_{lottery} \cdot V_{lottery}$ ). Similarly,  $EV_{surebet}$  denotes the expected value of surebet, which is simply the value of surebet here ( $EV_{surebet} = V_{surebet}$ , since  $P_{surebet} = 1$ ). In standard R formula syntax:

$$\text{choose\_lottery} \sim \text{delta\_EV} \quad (2.32)$$

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

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:

$$\text{choose\_lottery} \sim \text{delta\_ev} + \text{prev\_outcome} + (\text{delta\_ev} + \text{prev\_outcome}|\text{subjID}) \quad (2.33)$$

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

To test whether the  $\rho$  and  $\kappa$  estimates were different across species, we specified a linear model (LM) as follows:

$$\text{parameter} \sim \text{species} \quad (2.34)$$

where parameter is the MAP estimates of  $\rho$  or  $\kappa$ , species is a categorical variable with three levels: human, rat and mouse.

#### 2.2.4.4 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  $\rho$ , the prior over the utility exponent  $\rho$  was  $\text{Lognormal}(\log(0.9), 0.4)$ , a weakly informative prior that prefers  $\rho$  to be close to risk-neutral. For models containing  $\kappa$ , the prior was  $\mathcal{N}(0, 2)$ . The prior over noise  $\sigma$  was  $\text{Gamma}(6, 3)$ . The prior over the mixing weights  $\vec{\omega}$  was a Dirichlet distribution with the concentration parameter  $\alpha = [6, 2, 2]$ . The resulting  $\omega_{\text{rational}}$  distribution was broad and had the mean of 0.6, both  $\omega_{\text{lottery}}$  and  $\omega_{\text{surebet}}$  distribution had the mean of 0.2. By attributing more weight to the rational agent over the habitual agents, the prior reflected our selection of the experimental animals - only the ones with good psychometric curves were included. Four Markov chains with 1000 samples each were obtained for each model parameter after 1000 warm-up samples. The  $\hat{R}$  convergence diagnostic for each parameter was close to 1, indicating the chains mixed well.

#### 2.2.4.5 Synthetic datasets

To test the validity of our model, we first created synthetic datasets with parameters generated from the prior distributions described above. Each model was fit to its synthetic datasets and overall, they were able to recover the generative parameters accurately (FIGURE 3.15). This assured that our models had no systematic bias in estimating the parameters.

#### 2.2.4.6 Model prediction confidence intervals

To estimate the confidence intervals with model prediction as in FIGURE 2.3B, we first generated a synthetic dataset with out-of-sample lottery magnitudes (incremented by 1). After parameter sampling in each iteration (in the `generated quantities` block), the sampled parameters were used to predict the choices given the synthetic offers. The resulting output is a `n_iter` x `n_lott_mag` matrix, where `n_iter` is the number of iterations and `n_lott_mag` is the length of unique lottery magnitudes. Finally,  $\pm 80$ ,  $\pm 95$  and  $\pm 99$  confidence intervals for each offer were estimated by taking the respective quantiles of the `n_iter` predicted choices.

#### 2.2.4.7 Model comparison

To assess model performance on unseen data, we performed 10-fold cross-validation of the MCMC fits. For each fold, the model first fit on training set for that fold. We then computed (in the `generated quantities` block) the log predictive densities by passing in the held-out data, using the posterior 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

$$\widehat{lpd} = \sum_{i=1}^n \log \left( \frac{1}{S} \sum_{s=1}^S p(y_i | \theta^s) \right), \quad (2.35)$$

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

## 2.3 Results

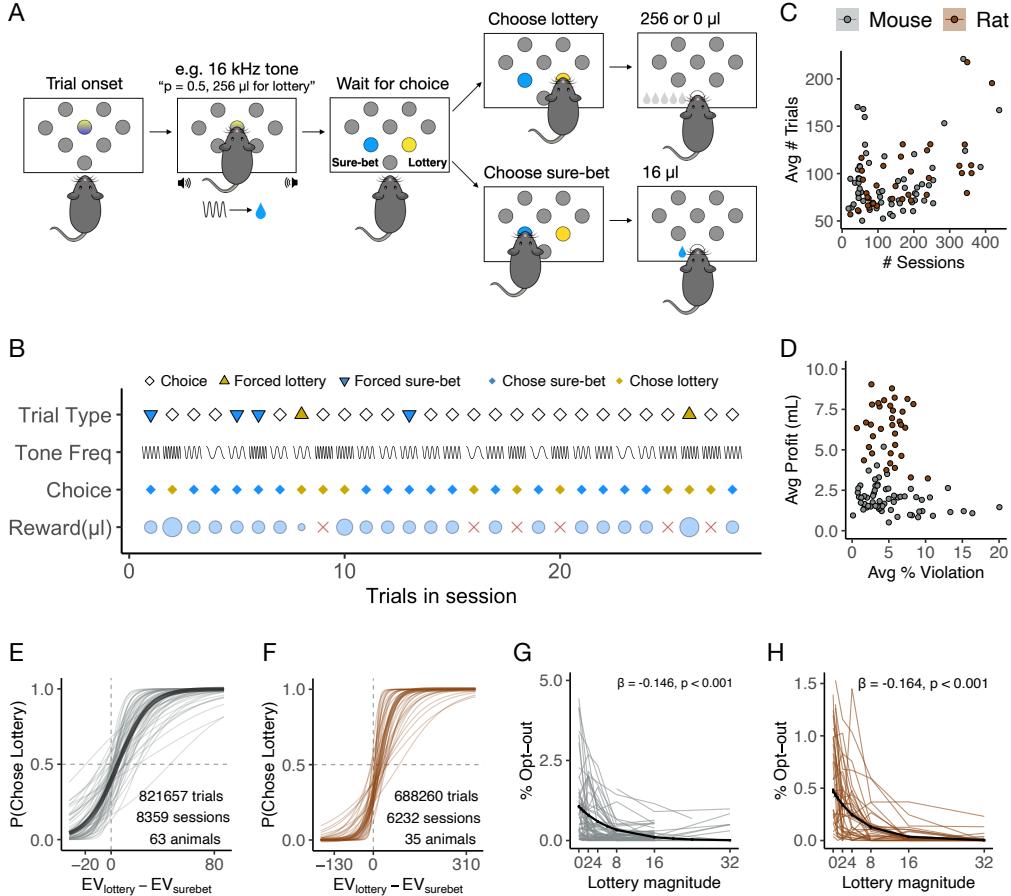
### 2.3.1 A cross-species paradigm of risky choice

Thirty-nine rats and sixty-four mice were trained on the risky choice task. On each trial, the animal chose between a surebet and a lottery with fixed probability and auditory-guided magnitude (FIGURE 2.1A-B, see task details in Methods). The data in this manuscript includes 14,591 behavioral sessions from rodents (8,359 from mice, 6,232 from rats), with a total of 1,509,917 choice trials (821,657 from mice, 688,260 from rats). On average, mice performed  $91.5 \pm 33.3$  (mean and standard deviation) choice trials per session, and a rat performed  $96.9 \pm 35.0$  choice trials per session (FIGURE 2.1C). A trial was considered a violation if the animal failed

to poke into central 300 s after trial start, or it did not make a choice 30 s after fixation, thereby ‘opting out’ this trial. The percentage of violation trials is a proxy of the animal’s productivity: the higher the violation rate, the fewer trials they completed. On average, mice opted-out  $5\% \pm 4$  and rat opted-out  $4.9\% \pm 2$  of all trials in a session, suggesting that both species were quite motivated to perform the task (FIGURE 2.1D).

The animals learned the task after extensive training. Rats required fewer training days than mice to approach the efficacy of an expected value maximizer than mice (mice:  $129 \pm 56$  days, rats:  $90 \pm 39$  days, FIGURE ??). 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 2.1E-F), some animals did not (e.g. FIGURE ?? 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 2.1G-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

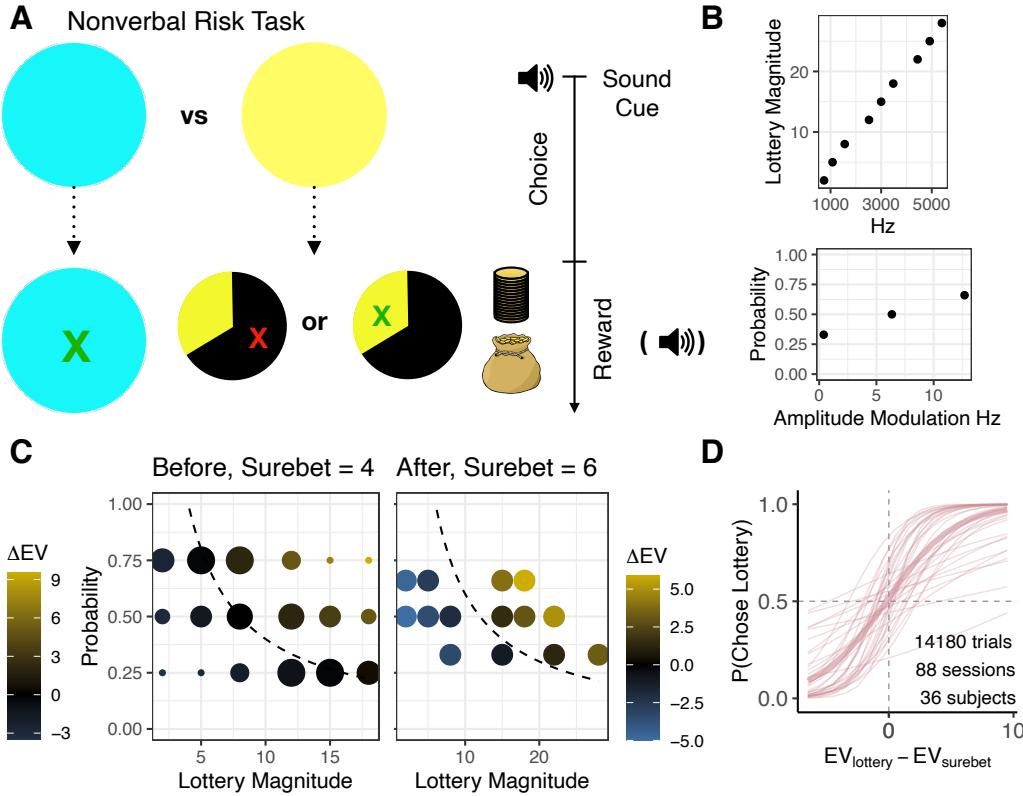


**Figure 2.1.** The risky choice task and animal behavior. **A.** Schematic of the risky choice task. **B.** An example trial sequence. For trial type, white diamond, yellow triangle and blue triangle represents the choice trial, forced lottery trial and forced surebet trial, respectively. The sine-waves 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 ( $\mu\text{l}$ ) 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 ( $V_{lottery}P_{lottery} - V_{surebet}$ ), where  $V$  represents  $\mu\text{L}$  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.  $\beta$  denotes the main effect of lottery magnitude on % opt-out. **H.** The opt-out behavior in rats.

between a surebet and a lottery option, whose magnitude and probability was signaled by a sound cue (FIGURE 2.2). 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).

### 2.3.2 Systematic modeling of risky choice

We considered three functional forms of risk. Firstly, we can model risk-tolerance  $\rho$  as the curvature of a power-law utility function,  $U = V^\rho$  (FIGURE 2.3A, 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 ( $\kappa$ , FIGURE 2.3A, BOTTOM), as in the mean-variance models.



**Figure 2.2.** The human 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 each repeated from  $n = 4$  to  $n = 16$ ). Right: choice set from the final version. Each offer (14 unique lotteries) was repeated exactly  $n = 8$  times. The dashed curve is the indifference line ( $ΔEV = 0$ ), 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 2.1E & F. Only now  $V$  represents number of coins.

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  $\sigma$  can represent noise in the utility function as in  $\mathcal{N} \sim (EU, \sigma)$ , where  $EU$  is the expected utility of lottery or surebet (FIGURE 2.3B, 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  $\mathcal{N} \sim (EU, \gamma EU)$  and  $\gamma$  represents the coefficient of variation (FIGURE 2.3B, BOTTOM).

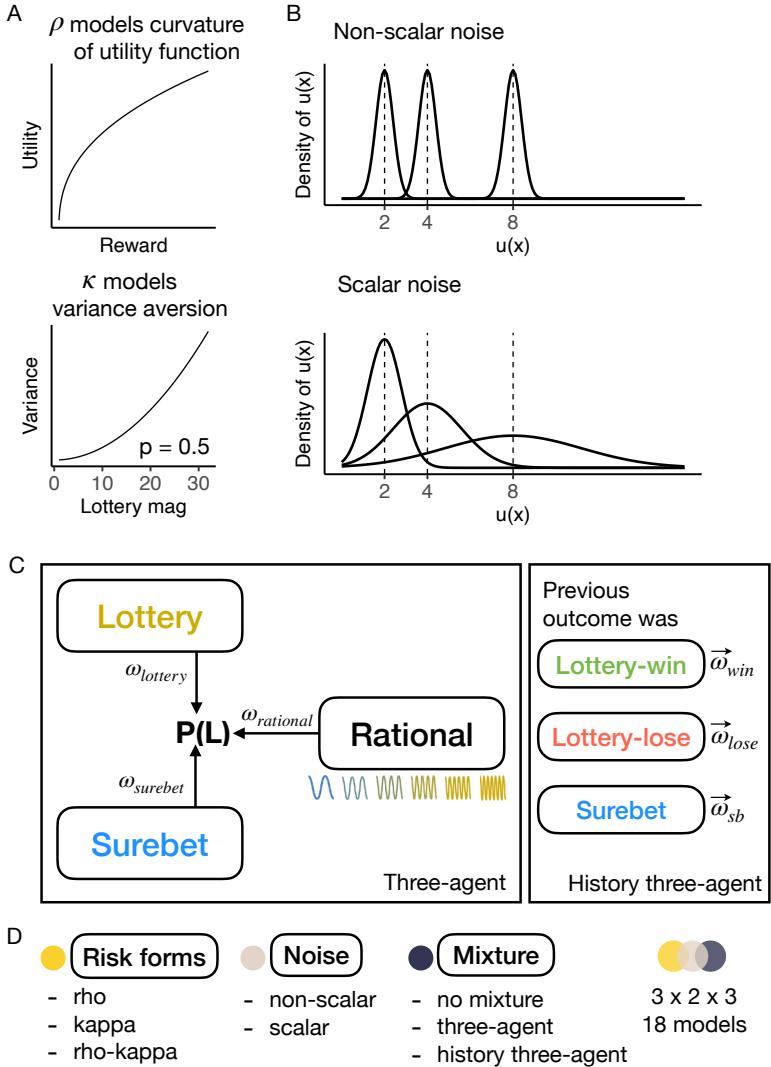
Finally, we considered three formulations of heuristic strategy mixtures. The first models subjects as a single utility-maximizing agent. Referred as the ‘rational’ agent, this agent can express any of combination of the risk forms and noise specifications. This is the standard approach in most human 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  $\omega$ , 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 2.3C, 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 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 2.3C, RIGHT).

With three risk forms, two noise specifications and three mixture formulations, we constructed 18 models of risky choice in total (FIGURE 2.3D, see all models in Methods). For naming conventions, ‘history-mix-rho-kappa-sigma’ refers to a model that uses both  $\rho$  and  $\kappa$  as functional forms of risk-tolerance, a standard  $\sigma$  noise specification and the history mixture formulation. Take another example, ‘rho-scalar’ refers to a model using  $\rho$  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 2.9). We estimated the joint posterior over the parameters for each subject separately using Hamiltonian Monte Carlo sampling in Stan (Carpenter et al., 2016).

### 2.3.3 Model-based estimation of risk attitudes

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 2.4B, more examples in FIGURE 2.10, 2.11, 2.12). 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 2.4A LEFT). Similarly, all rats were also estimated to be risk-averse by



**Figure 2.3.** Systematic modeling of risky choice. **A.** Functional forms of risk. Top:  $\rho$  models the curvature of utility function as in standard economic models. Bottom:  $\kappa$  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  $\omega$ , 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.

this model ( $0.61[0.37, 0.96]$ , FIGURE 2.4A 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 2.4A RIGHT). As can be observed from the distribution of posterior  $\rho$  samples (FIGURE 2.4C, 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  $\rho$  estimates, see Methods for details).

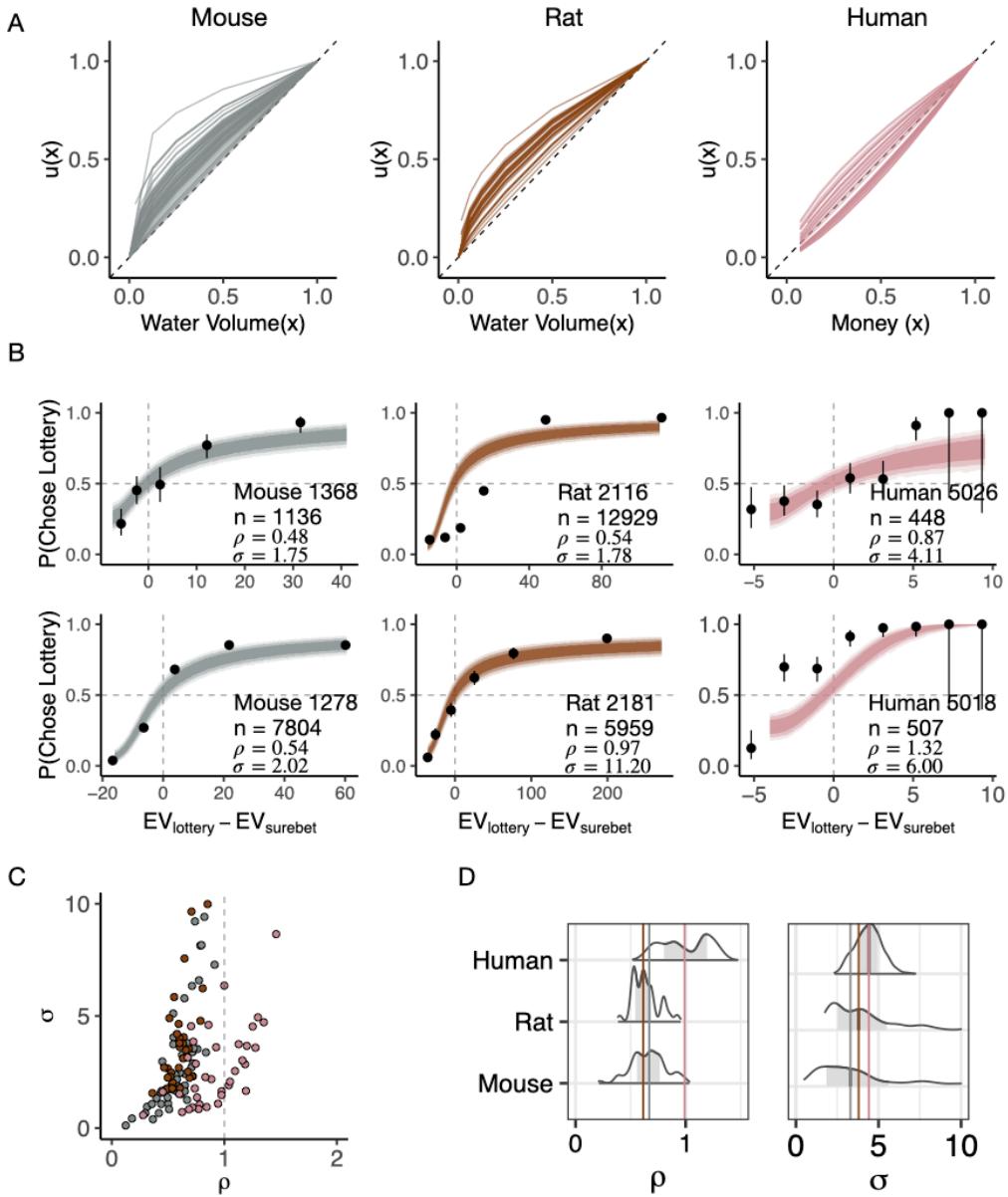
The concave utility functions found in rats are consistent with previous results (Constantinople et al., 2019b). To our knowledge, this is the first dataset describing the curvature of utility functions in mice. We found that mice and rats did not differ significantly (when parameterized by the rho-sigma 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).

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 2.4B). As a comparison, we examined the estimates of  $\rho$  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  $\rho$  shifted substantially (FIGURE 2.5A). The median  $\rho$  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  $\rho$  was due to the addition of variance-aversion,  $\kappa$ , which, like  $\rho$ , influences the indif-

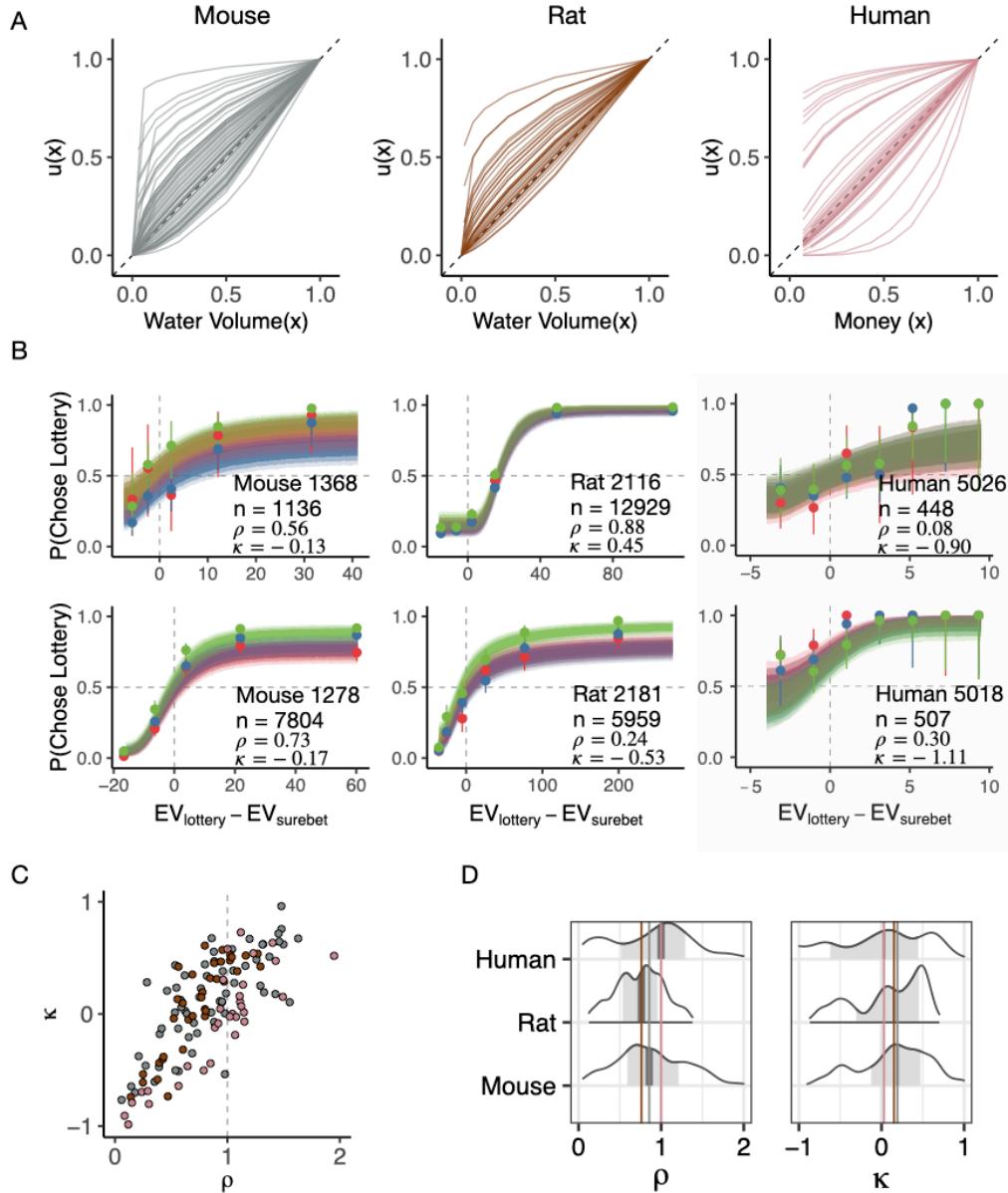
ference between lotteries and the surebet ( $\kappa$  median & 95%CI: mice  $0.20[-0.72, 0.97]$ , rats  $0.19[-0.72, 0.92]$ , humans  $0.01[-1.17, 0.78]$ ; FIGURE 2.5C, D). In addition, the heuristic mixture formulation enabled the model to capture stimulus-independent biases, fitting subjects like 2116 better while also more accurately estimating its utility function (FIGURE 2.5B). Moreover, the incorporation of trial-history-dependent parameters allowed the history effects in subjects like 1368, 1278 and 2181 to be estimated (FIGURE 2.5B). On the population level, the MAP  $\rho$  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  $\kappa$  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 2.5C,D). Overall, these results suggest that once we isolated the history-dependence and stimulus-independent biases, the distribution of  $\rho$  and  $\kappa$  across species become more overlapped.

### 2.3.4 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 2.19,



**Figure 2.4.** Fits by the rho-sigma model. **A.** The subjective utility functions for each subject computed using the maximum a posteriori (MAP)  $\rho$  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  $\rho$  and  $\sigma$  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 2.5.** Fits by the history-mix-rho-kappa-sigma model. **A.** Same as FIGURE 2.4A. **B.** Same as FIGURE 2.4B 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  $\rho$  and  $\kappa$  MAP estimates. **D.** Same as FIGURE 2.4D.

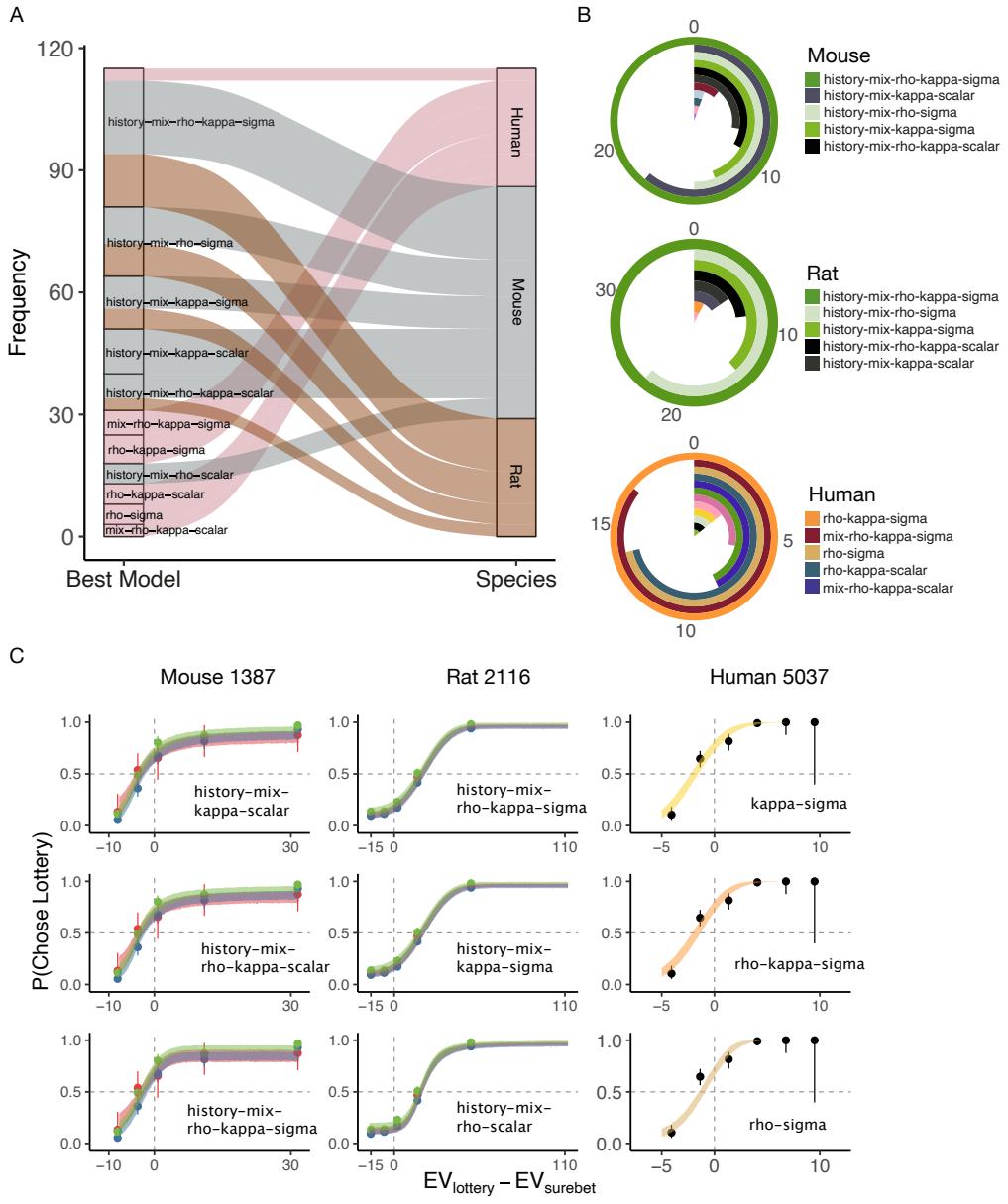
2.20, 2.21), differences emerged between humans and rodents on the population level. For both mice and rats, the most frequent best-fitting model was ‘history-mix-rho-kappa-sigma’ (best model in 28.6% mice, 37.1% rats), suggesting that rodents relied on heuristics as part of their strategy and were sensitive to the previous trial’s outcomes, consistent with previous reports (FIGURE 2.6A, B). Their noise was best captured by a simple  $\sigma$  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 2.6B). 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 2.19 1108, 1121, 1297, 2.20 2105, 2119, 2136), and several models performance was virtually indistinguishable by eye (FIGURE 2.6C). 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 2.6A, 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 2.21), performance of the reported ‘best-fitting’ models were in fact indistinguishable from many other models FIGURE 2.6C). 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.

### 2.3.5 Estimation of rho is model-dependent

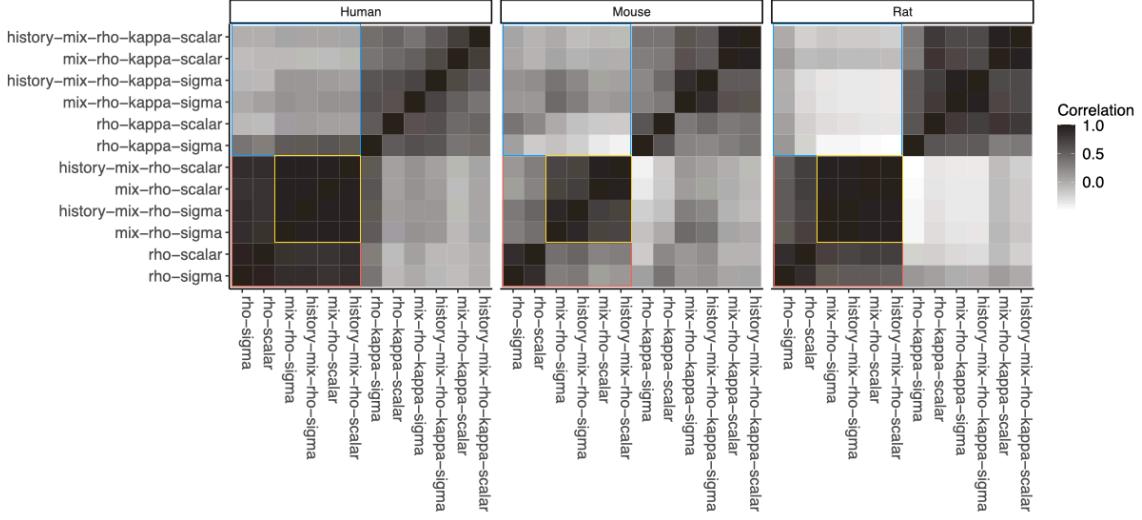
Our model comparison analysis found that some classes of models were better than others while failing to definitively identify a best model for each subject. Can we still learn something from the estimation of risk attitudes from different models? To compare the estimation of  $\rho$  from all the models containing  $\rho$  ( $n = 12$ ), we first extracted the MAP  $\rho$  estimate from each model for each individual, and then performed a rank correlation test on the model estimates for each pair of  $\rho$ -containing models. The resulting correlation matrices revealed interesting patterns within and between species (FIGURE 2.7). The most conspicuous finding was that  $\rho$  estimates from models fitting  $\rho$  and  $\kappa$  simultaneously had low to negative correlations with  $\rho$  estimates from models only containing  $\rho$  (FIGURE 2.7, 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  $\rho$  or  $\kappa$  (but not both), that parameter strongly influences the indifference point of the psychometric curve. When



**Figure 2.6.** 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 2.12. **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.

both parameters are included, they trade off (FIGURE 2.5C).

The next finding was that while a high correlation of  $\rho$  was observed within models only estimating  $\rho$  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 2.7, HUMAN, 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  $\rho$  across models with and without these elements is low ( $\rho$  (FIGURE 2.7, MOUSE AND RAT, YELLOW SQUARE VS. RED SQUARE)), we infer that the inclusion of these factors is important for correctly estimating  $\rho$ . 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  $\rho$  estimation when other model properties were matched. For example,  $\rho$  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,  $\rho$  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  $\rho$  depends on the model as well as behavioral tendencies in the dataset.



**Figure 2.7.** Species correlation of  $\rho$  estimates from all the models containing  $\rho$ . MAP  $\rho$  estimate from each model was obtained for each individual. Spearman’s rank correlation test was performed between each pairwise model’s estimation.

## 2.4 Discussion

Risky choice has been widely researched in humans and non-human animals. However, contextual 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  $\rho$  and  $\kappa$  became more overlapping across species, and each population was

comprised of risk-averse and risk-seeking individuals. Moreover, we showed that the estimation of  $\rho$  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 2.5A). 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  $\rho$  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, ?), 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  $\kappa$ , 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,

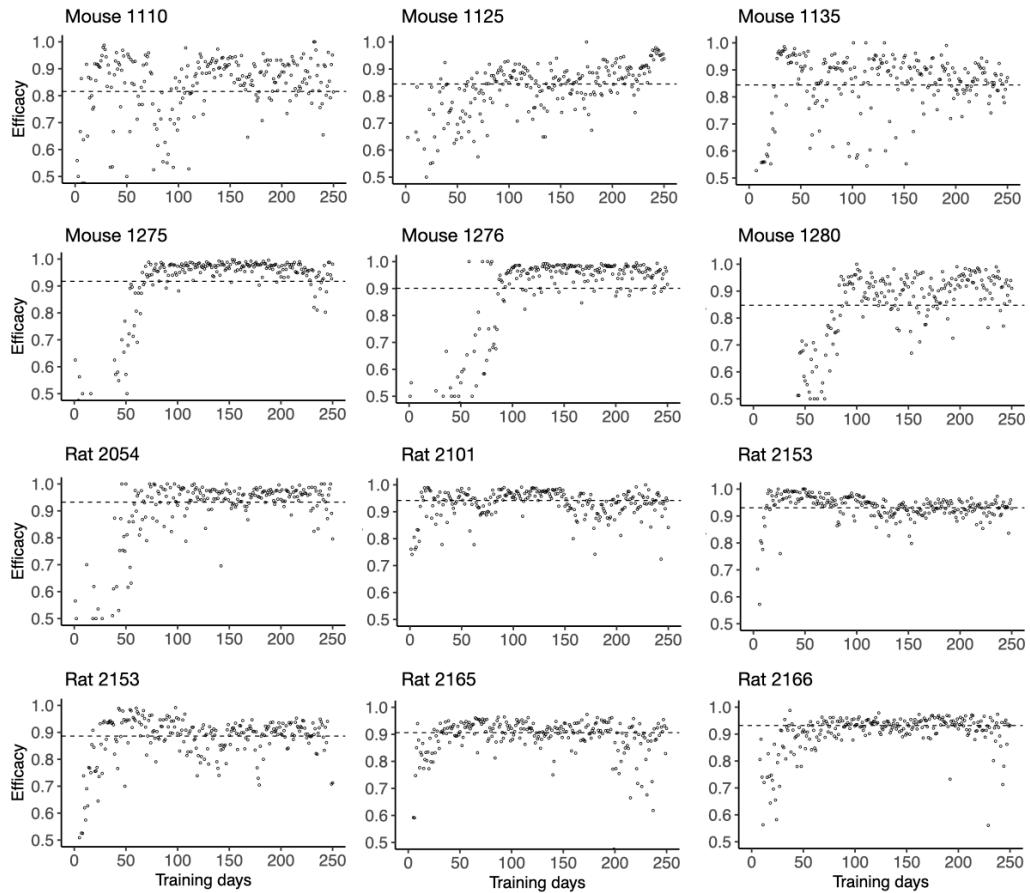
Balci et al., 2009, Hanks and Summerfield, 2017, de Gee et al., 2018), studies that address the description-experience gap in risky choice abound (Hertwig and Erev, 2009). Increased risk-seeking was found for experienced gambles of equally probable outcomes relative to described ones, and importantly, 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 in-

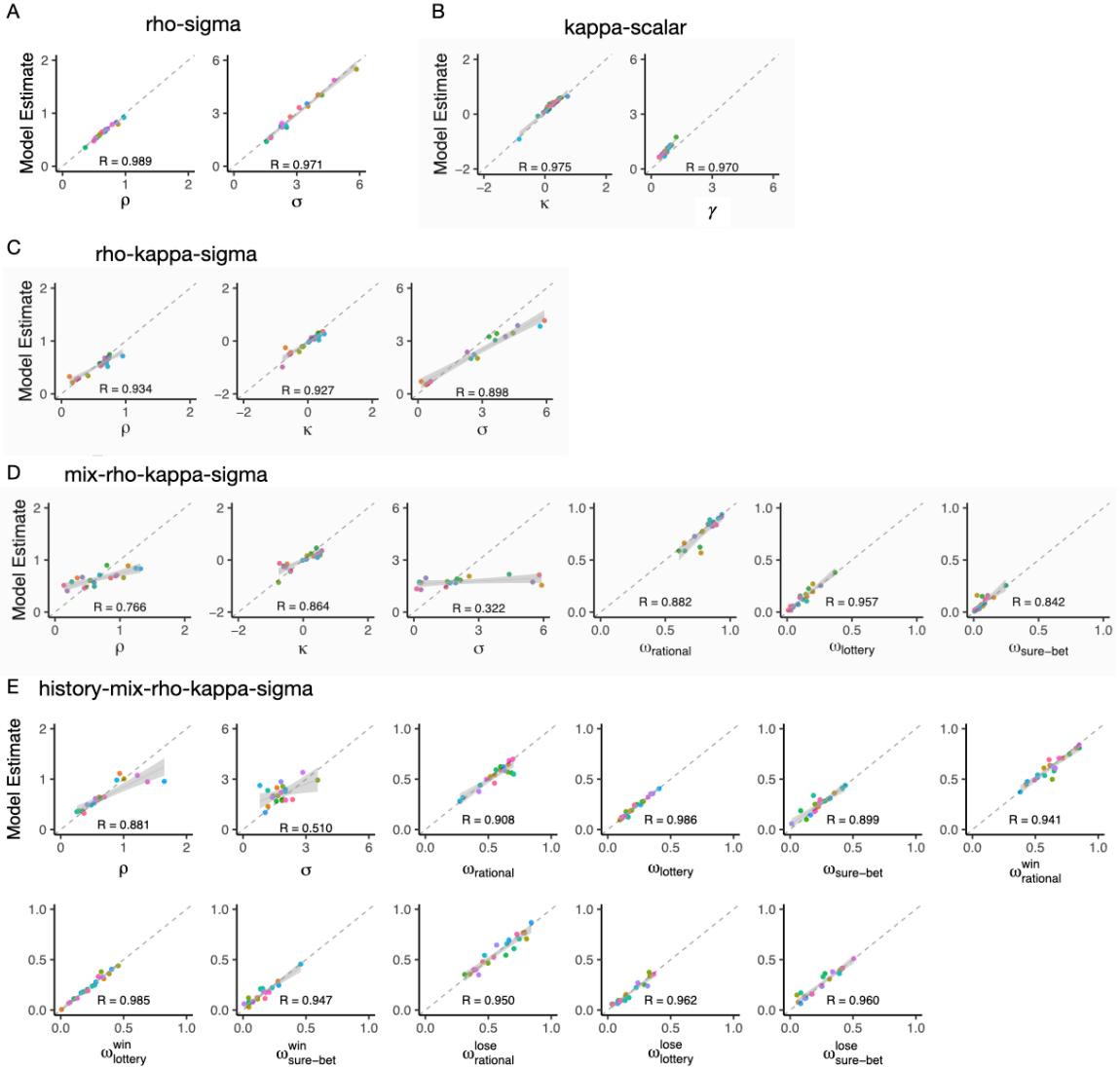
structured learning and experience. The finding that some humans (e.g. FIGURE 2.21 5064, 5059, 5100) were best fit by a mixture model suggests the use of heuristics. Robust heuristics rather than optimal inferences are often adopted under uncertainty, which was probably what the task appeared to these 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.

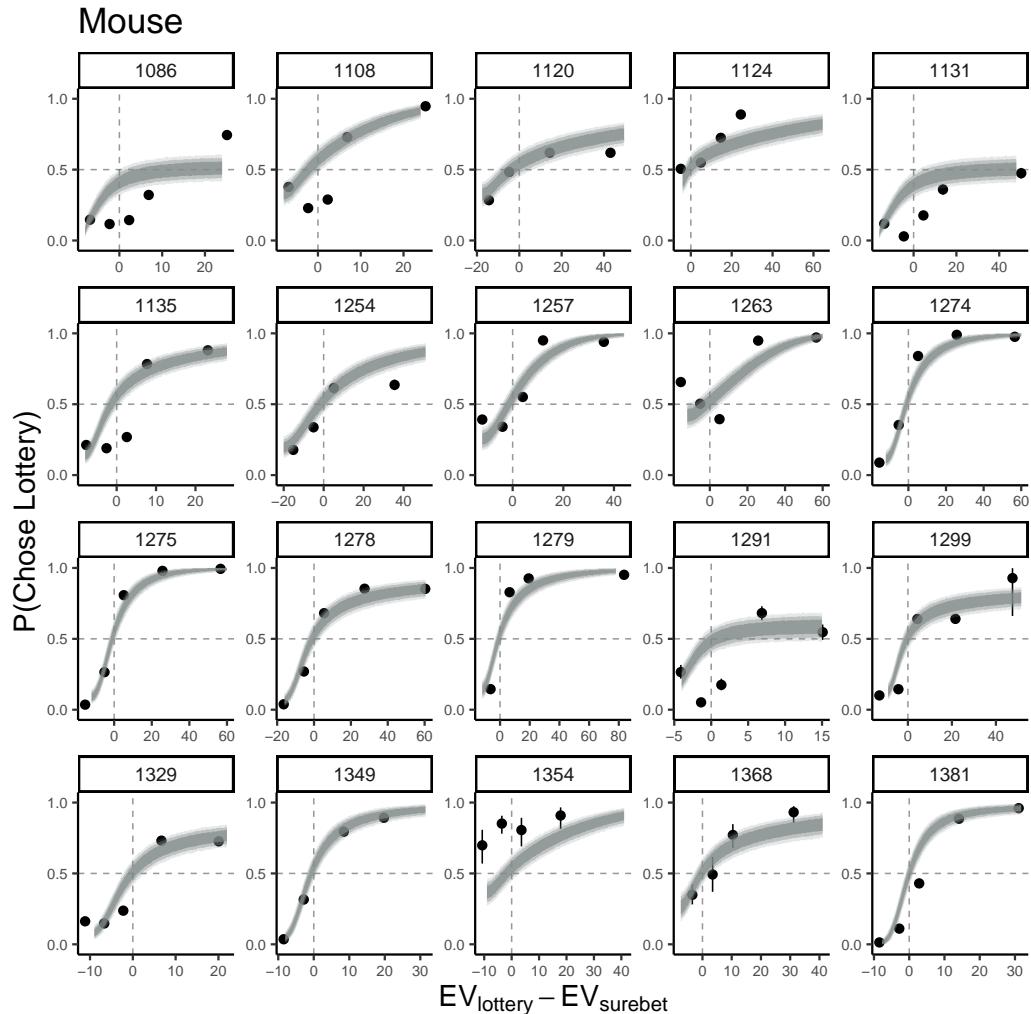
## 2.5 Supplementary figures



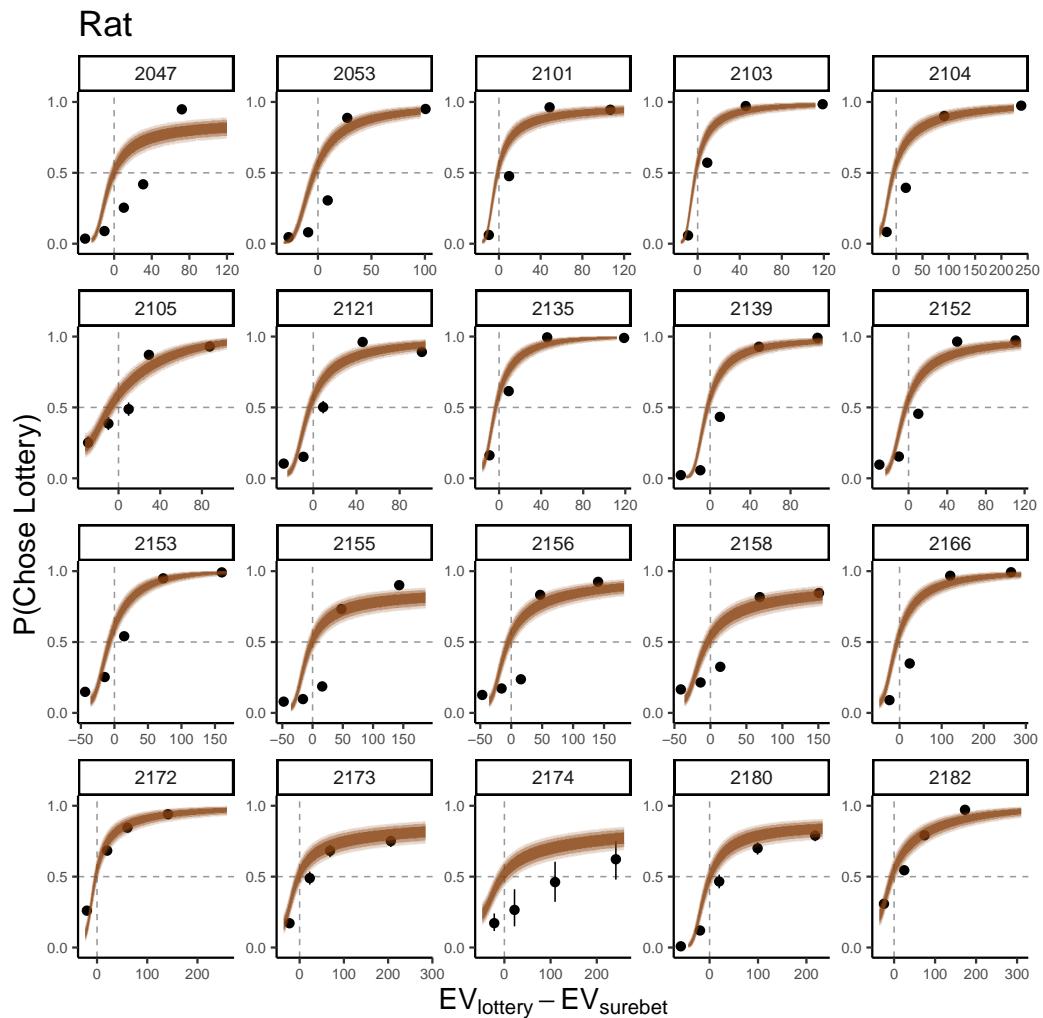
**Figure 2.8.** Example behavioral performance over training. Animal's 'efficacy' on each training session was computed by comparing its choices against that from a random agent and from a perfect expected value maximizer. See Methods for details. The dashed line indicates the median efficacy over 250 training days.



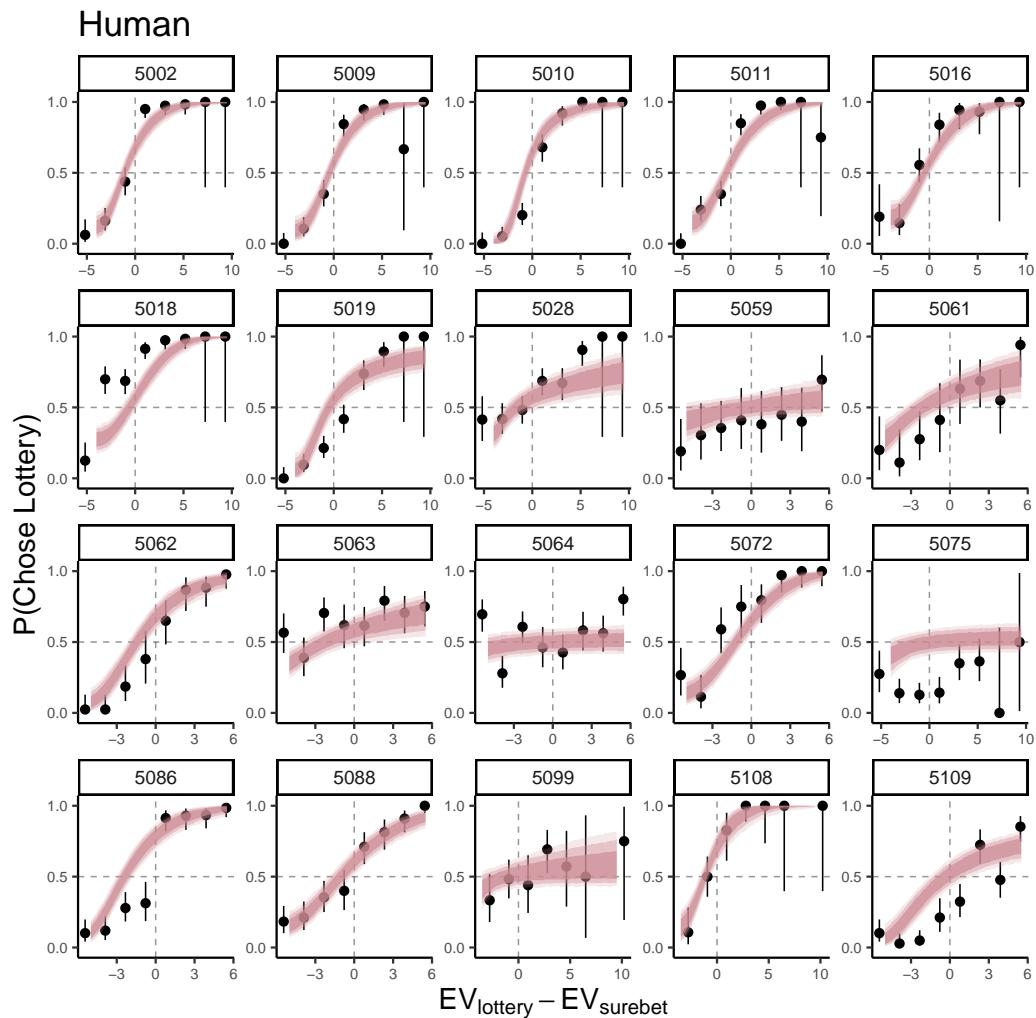
**Figure 2.9.** The models can recover the data-generating parameters well. Twenty Synthetic datasets were created by sampling from aforementioned prior distributions. The true parameter values are on the x-axis, MAP model estimates are on the y-axis. Color represents the identity of each synthetic dataset. Only 5 out of all 18 models are shown here.



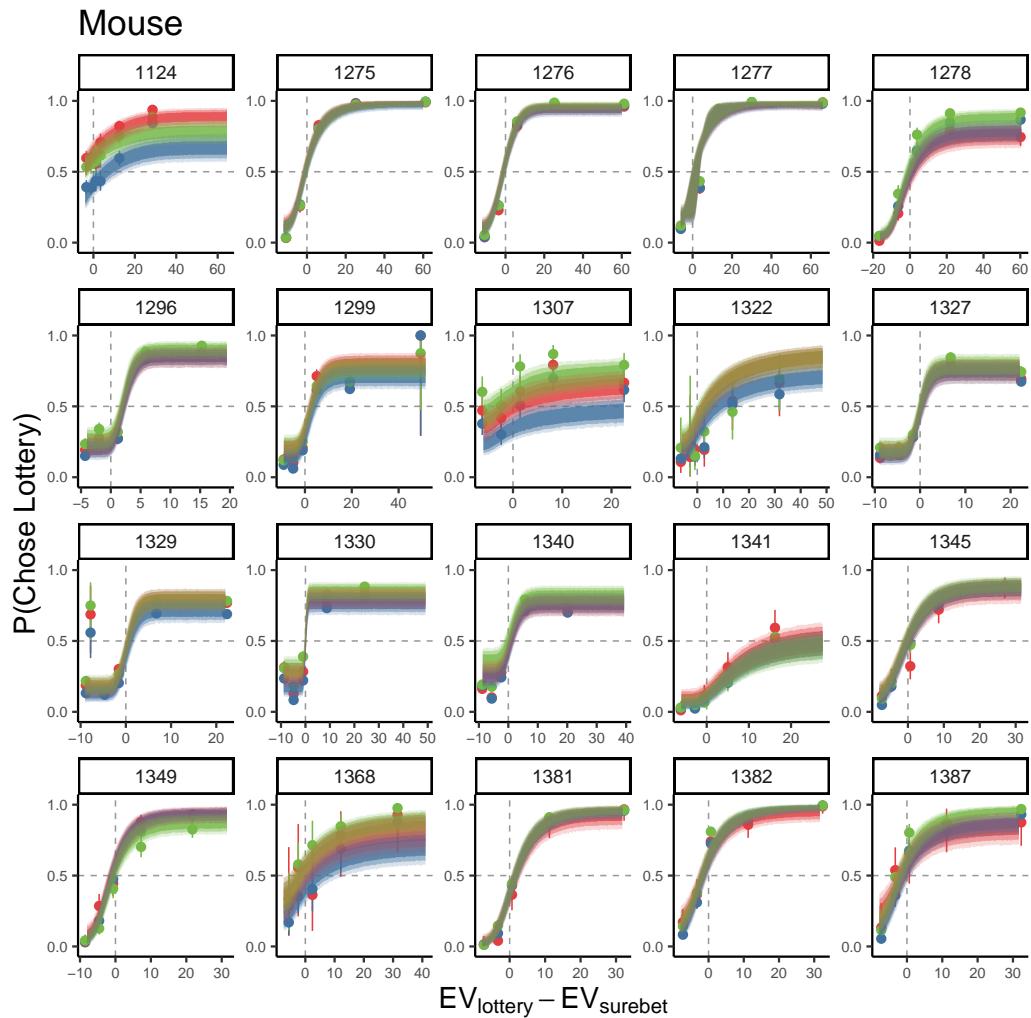
**Figure 2.10.** Example mouse's choices superimposed with predictions from the rho-sigma model. The circles with error bars are the binned mean and 95% binomial confidence intervals. The ribbons are model predictions generated using fitted parameters from the rho-sigma model. The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively.



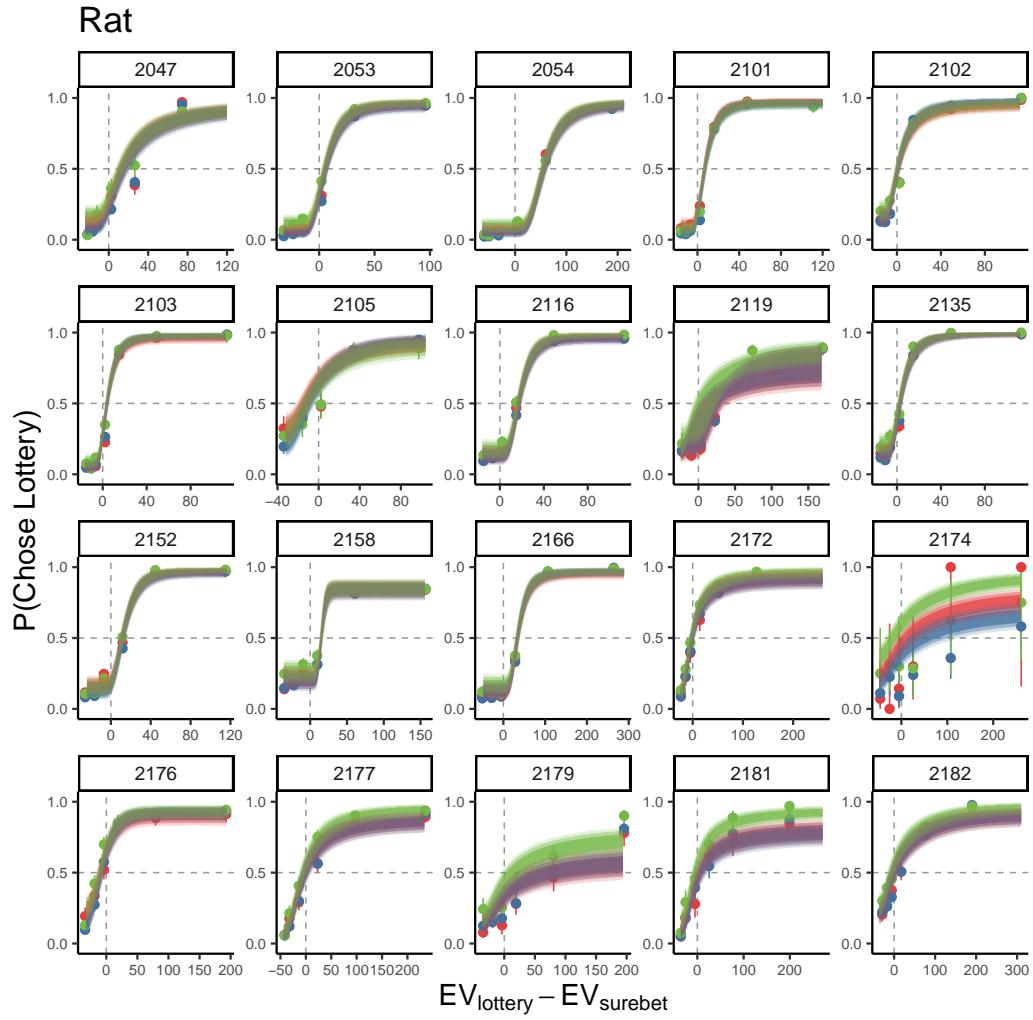
**Figure 2.11.** Example rat's choices superimposed with predictions from the rho-sigma model.



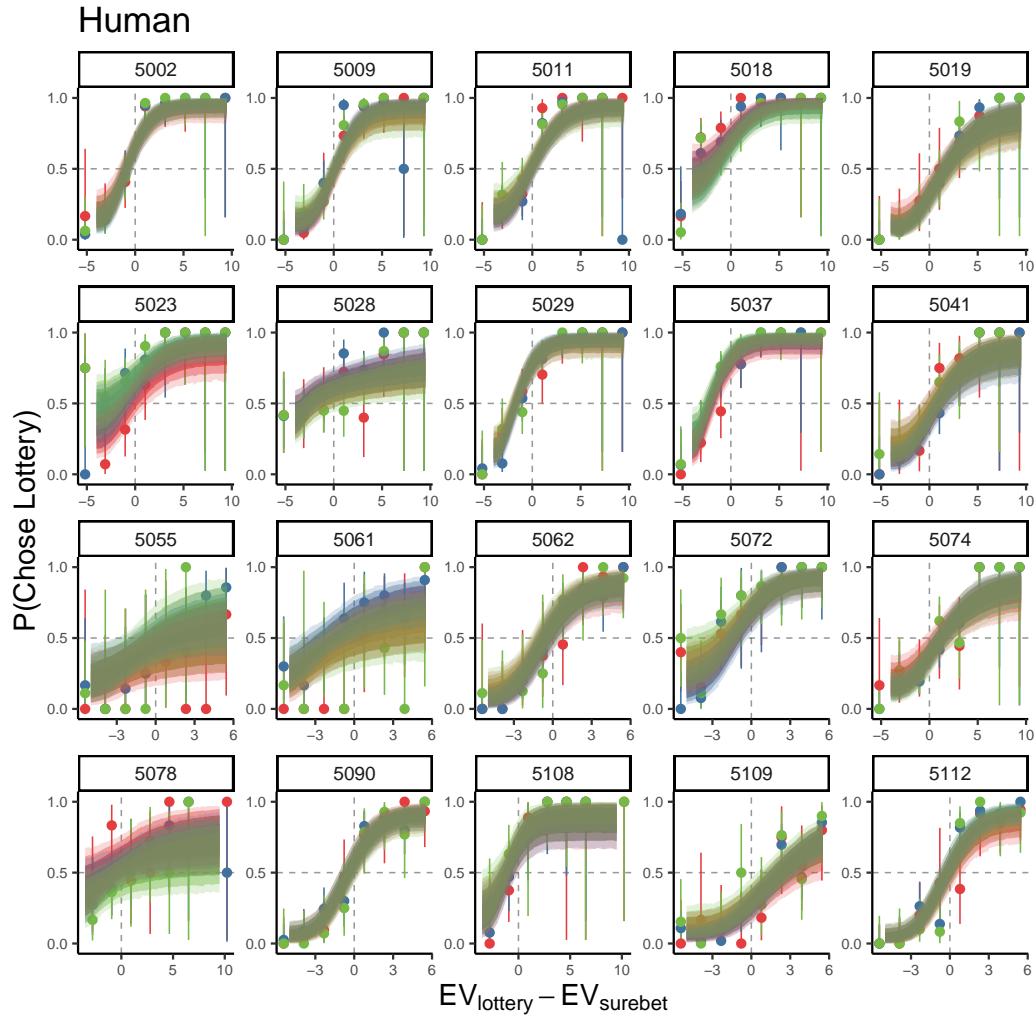
**Figure 2.12.** Example binned human's choices superimposed with predictions from the rho-sigma model.



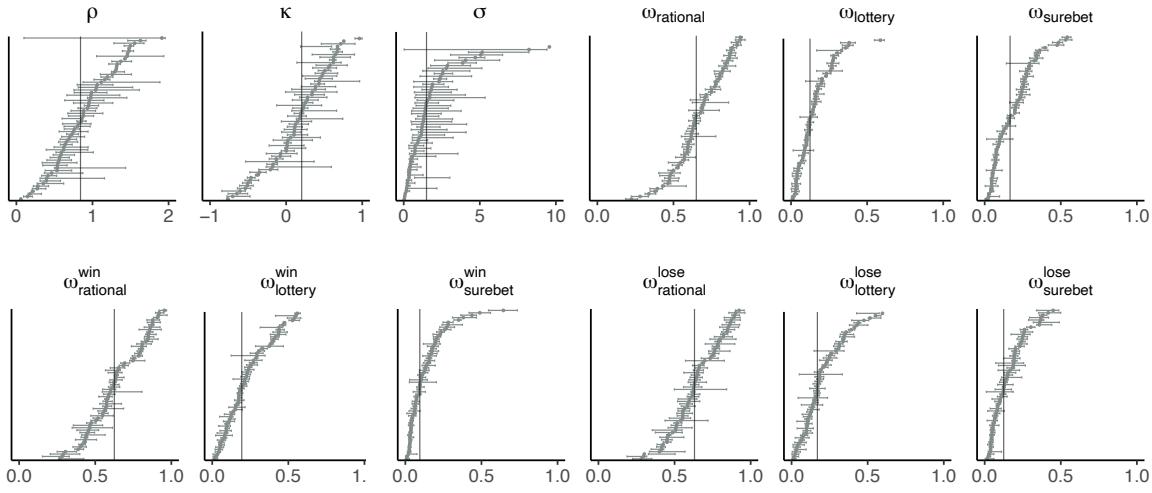
**Figure 2.13.** Example mouse's choices superimposed with predictions from the history-mix-rho-kappa-sigma model. The color represents the previous trial's outcome: lottery-win = green, lottery-lose = red, surebet = blue.



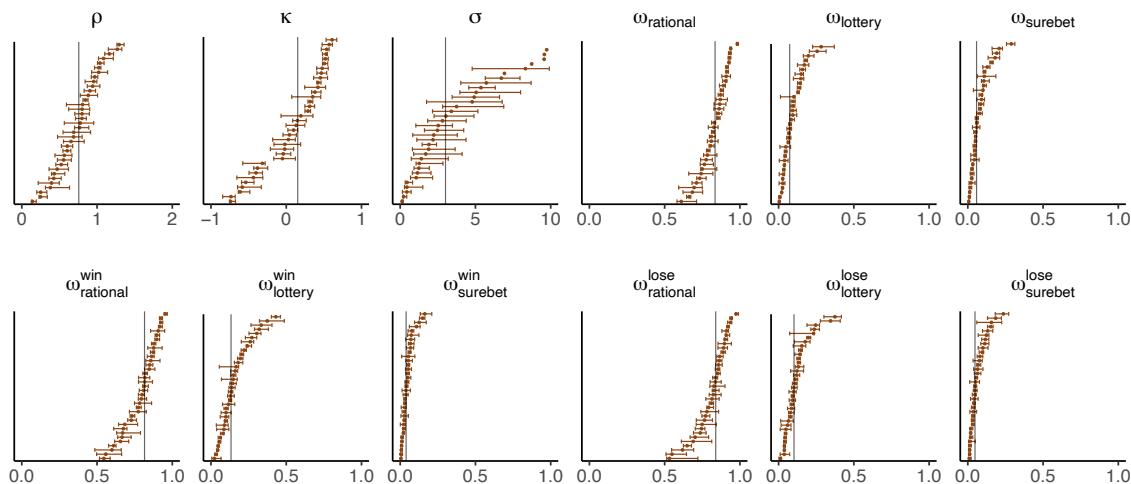
**Figure 2.14.** Example rat's choices superimposed with predictions from the history-mix-rho-kappa-sigma model.



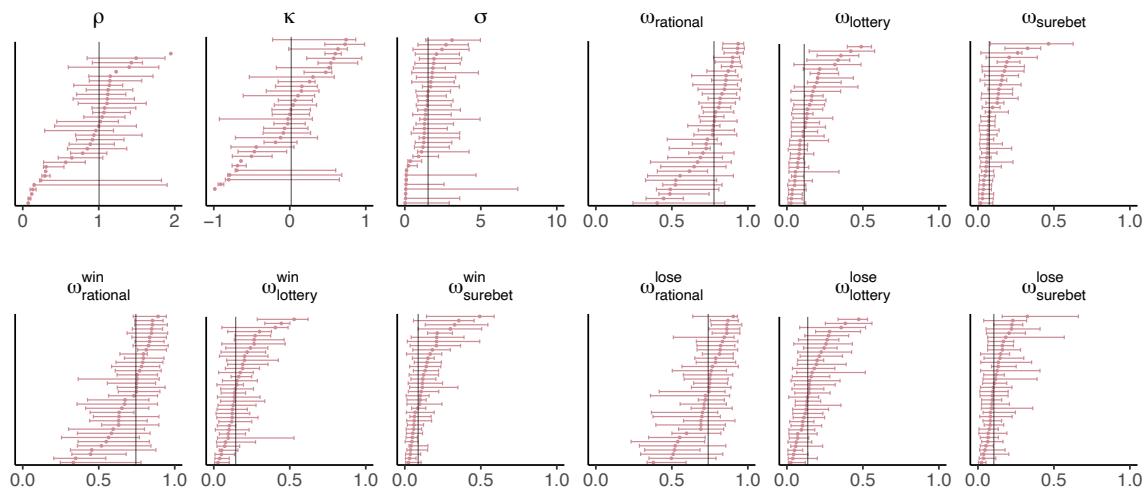
**Figure 2.15.** Example binned human's choices superimposed with predictions from the history-mix-rho-kappa-sigma model.



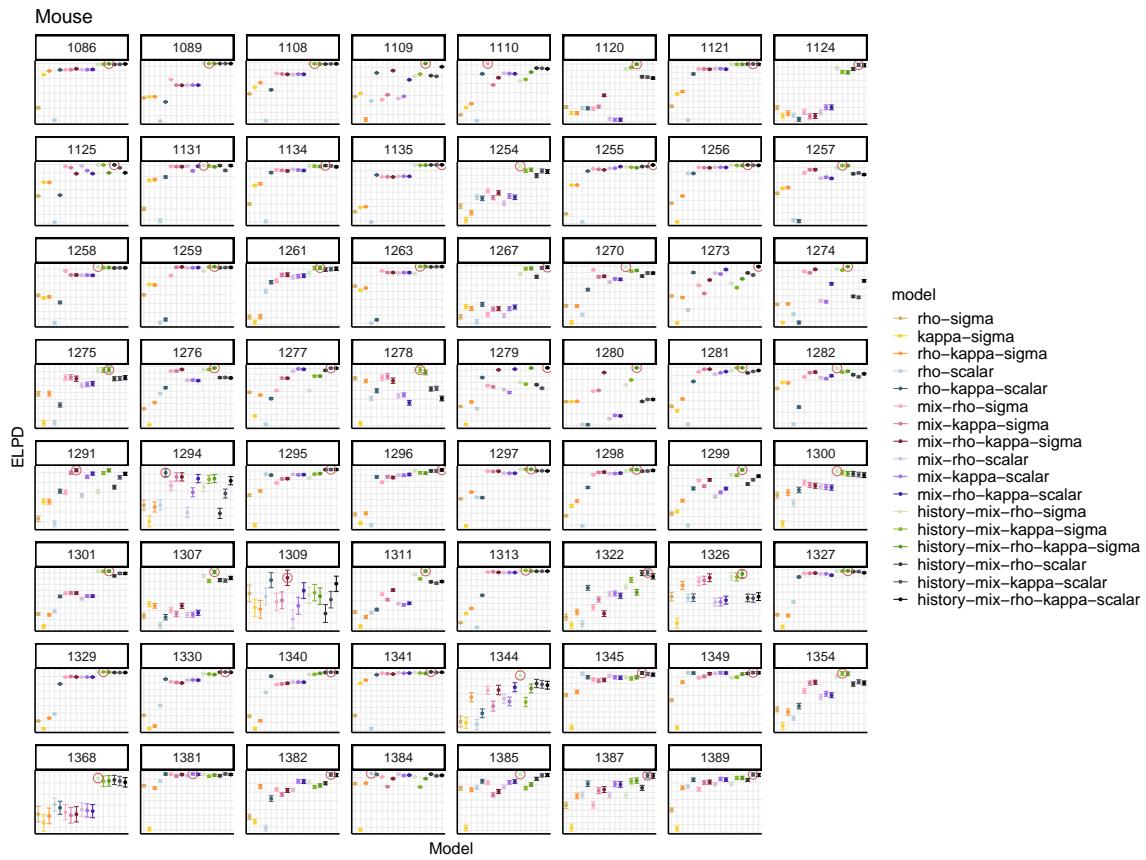
**Figure 2.16.** MAP parameter estimation with 95 % C.I. fit by the history-mixture-rho-kappa-sigma model for each mouse. The black bar indicates the median of concatenated posteriori samples across all subjects



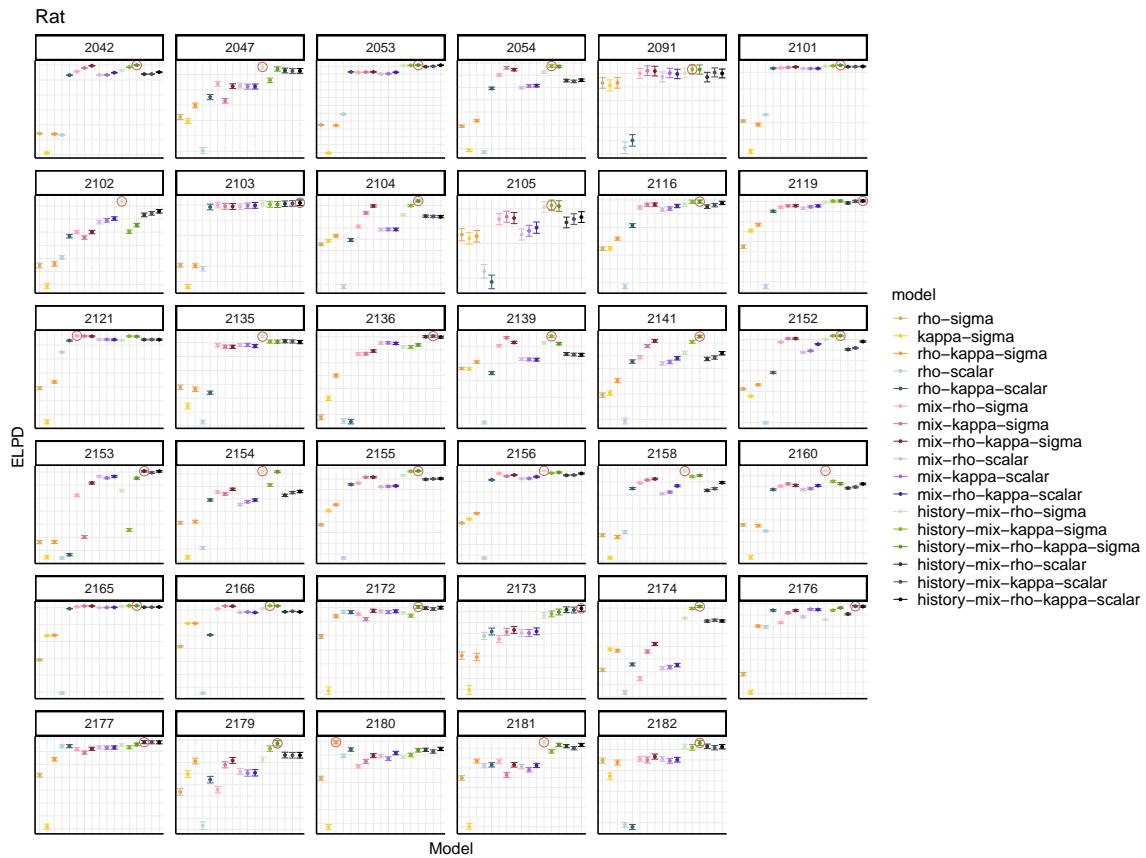
**Figure 2.17.** MAP parameter estimation with 95 % C.I. fit by the history-mixture-rho-kappa-sigma model for each rat.



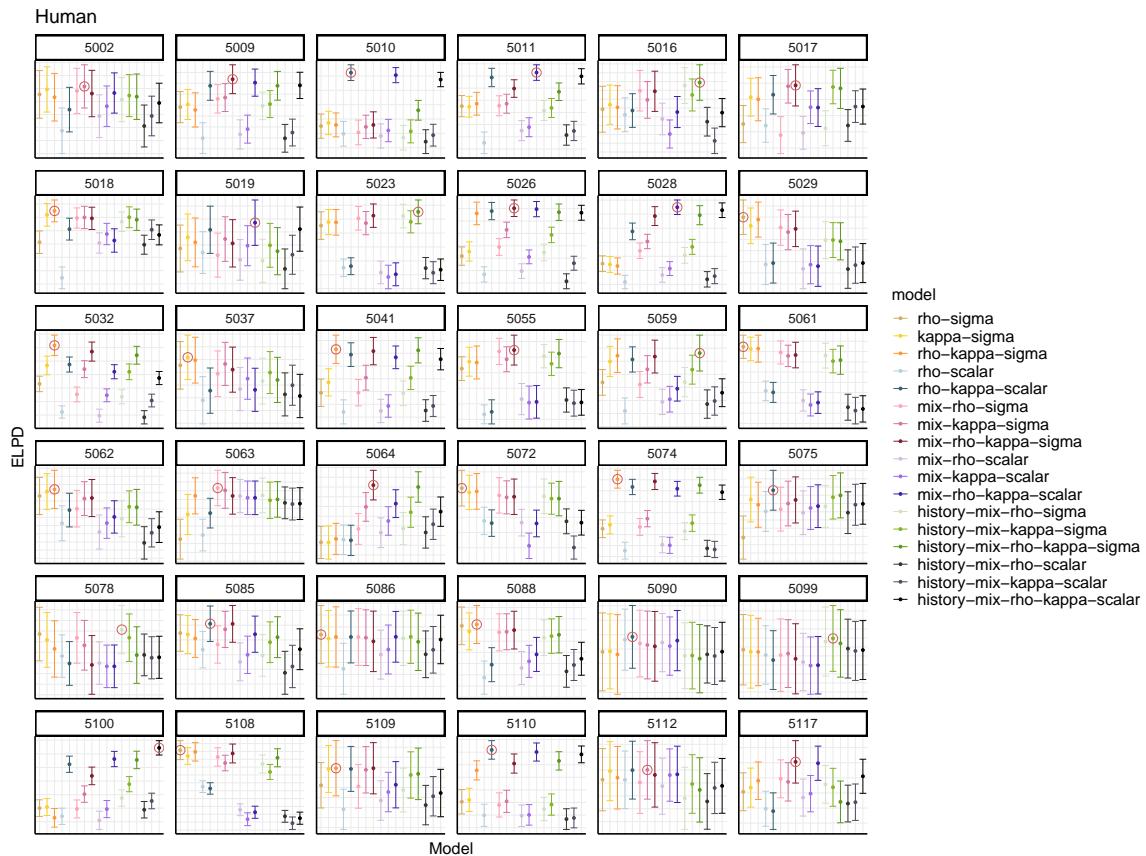
**Figure 2.18.** MAP parameter estimation with 95 % C.I. fit by the history-mixture-rho-kappa-sigma model for each human subject.



**Figure 2.19.** 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 2.20.** 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 2.21.** 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.

# **Chapter 3**

## **Frontal but not parietal cortex is required for economic decisions under risk**

### **3.1 Introduction**

Understanding 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 pathologies (Clifton et al., 2018). On the other hand, inadequate risk-taking is also undesirable: 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 (23and Me Research Team et al., 2019, Xuan et al., 2017, Rao et al., 2018, Anokhin et al., 2009) suggest that ge-

netics accounts for a moderate proportion ( $\sim 30\%$ ) of variation in risk-tolerance, indicating that animal models can help establish the link between genes, brains and risk-tolerance.

The neurobiology of risky decision-making has been studied in human, non-human primate and rodent subjects (reviewed in Padoa-Schioppa, 2011, Orsini et al., 2015, Platt and Huettel, 2008). Rodent work has mostly focused on decision-making under uncertainty in new or changing environments: that is, the neural mechanisms for learning the values of actions. Action-value learning under ‘unexpected uncertainty’ has also been studied extensively in monkeys and humans (O’Doherty et al., 2015, Behrens et al., 2007, Monosov et al., 2015). These studies typically identify regions that are associated with learning action-values in general: amygdala (Orsini et al., 2017, Saez et al., 2017, Larkin et al., 2016), basal ganglia (Samejima et al., 2005, Stopper et al., 2013, Floresco et al., 2018), and orbital and medial prefrontal cortex (Stopper et al., 2012, Ogawa et al., 2013, van Holstein and Floresco, 2020). Human, monkey and, to a lesser extent, rodent work has also examined the neurobiology of decision-making under risk when the probabilities are known, i.e. ‘expected uncertainty’ (Kobayashi and Hsu, 2017, Christopoulos et al., 2009, Constantinople et al., 2019b, Hocker et al., 2021). This is closer to the way risky decisions are studied in economics or finance research, where the potential outcomes of different actions are given explicitly on each trial. In these studies, activity in regions associated with orienting decisions (Hanks et al., 2015), including the parietal cortex (Platt and Glimcher, 1999, Dorris and Glimcher, 2004) and frontal cortex (Chen and Stuphorn, 2015, 2018) represents the value of the options, as the subjects were typically asked to respond by shifting gaze to a spatial target.

One challenge in synthesizing the vast literature on decisions under risk is that risk-tolerance is not monolithic (Yates, 1992). When behavior is measured either in the laboratory or in real-life, any avoidance of uncertainty can be considered as ‘risk-aversion’, but such avoidance can come from distinct cognitive constructs. For example, risky behavior in teenagers may result from an incomplete perception of risk associated with those behaviors rather than a greater tolerance for the actual risk (Cohn et al., 1995). Reinforcement learning agents with high learning rates will seem more risk-averse because they will avoid actions after a single loss (i.e. ‘lose-shift’), even if on average that action provides good outcomes (March, 1996, Niv et al., 2002). Under the expected utility framework, risk-aversion is usually associated with a decelerating utility function: the more rapid the deceleration, the more risk-averse the subject (Von Neumann and Morgenstern, 1953, Rabin and Thaler, 2001, Kahneman and Tversky, 1979). In finance, risk-aversion is typically modeled as variance-aversion (Markowitz, 1968). This rich taxonomy of constructs underlying risk-preference not only adds confusion when parsing the literature, but also makes the design of animal experiments estimating all elements simultaneously difficult.

Here, we present results from a risky choice task where the animals make choices under ‘expected uncertainty’ on a trial-by-trial basis. On each trial, the rat makes an informed decision between a ‘surebet’ (small but guaranteed reward) and a lottery with fixed probability and cue-guided magnitude. Our model-based quantification of animals’ behavior incorporates parameters to capture marginal utility, decision noise, and choice biases. This task and modeling framework provides a foundation for rigorous circuit level exploration of the neurobiology of risky choice. With this framework, we examined the causal contribution of the frontal orienting field (FOF)

in frontal cortex and the posterior parietal cortex (PPC): two cortical areas that have been implicated in perceptual decision-making (Erlich et al., 2015, Hanks et al., 2015, Zhong et al., 2019, Raposo et al., 2014).

In perceptual decision that require working-memory, the FOF seems to be essential for maintaining a plan of the upcoming choice. Unilateral silencing of FOF biases animals towards the ipsilateral choice and this bias is larger for trials with longer memory periods (Piet et al., 2017). Bilateral silencing also generates an impairment that grows with longer delays or periods of integration (Erlich et al., 2015). In order to distinguish the role of working-memory from the cognitive processes required for economic choice under risk (i.e. trading off the cost of uncertainty with the benefits of a larger reward), we did not include a working-memory component in our task. Nonetheless, we predicted that unilateral silencing of FOF would cause contralateral impairments in economic choices, due to our hypothesis that FOF serves as a bottleneck for higher order cognitive processes to guide orienting decisions. For bilateral FOF the prediction was less clear: we expected it to influence behavior, possibly by increasing the decision noise.

In contrast to the results from FOF, we previously found that unilateral silencing of PPC in rats did not bias perceptual decisions. PPC only biased ‘free-choice’ trials, where the animal was rewarded regardless of the left or right response (Erlich et al., 2015). We speculated that the difference between the efficacy of PPC inactivations in perceptual vs. free choice might be that PPC only plays a causal role when decisions are internally guided. Risky choices are internally guided in the sense that each subject has some risk-preference: there is no single ‘correct answer’ on each trial. Moreover, signatures of expected value are reliably found in PPC (Dorris and Glim-

cher, 2004, Platt and Glimcher, 1999). Thus, we hypothesized that PPC silencing might influence economic choices, in contrast to perceptual decisions.

Although there is substantial literature comparing different functional forms of decision under risk (Farashahi et al., 2019, Spitmaan et al., 2019, Heilbronner, 2017), we are unaware of any previous studies that simultaneously estimates these parameters, examines how silencing of frontal or parietal cortices shifts specific cognitive constructs underlying risky choice. We found that PPC silencing had minimal effects on decisions under risk (but biased free choice). Surprisingly, we found that bilateral silencing of FOF shifted animals away from choosing the lottery. Model-based analysis of these results indicated that the shift was likely caused by a change in the utility function. Moreover, this effect can be parsimoniously explained by a dynamical model where the FOF part of a network for encoding the value of the lottery.

## 3.2 Materials and methods

### 3.2.1 Subjects

Six male Sprague Dawley rats and two male Brown Norway rats (Vital River, Beijing, China) were used in this study. Rats 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. Rats 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.

### **3.2.2 Behavior**

#### **3.2.2.1 Behavioral Apparatus**

Animal training took place in custom behavioral chambers, located inside sound- and light-attenuated boxes. Each chamber ( $23 \times 23 \times 23$  cm) was fit with 8 nose ports arranged in four rows (FIGURE 3.1A), with speakers located on the left and right side. Each nose port contained a pair of blue and a pair of yellow light emitting diodes (LED) for delivering visual stimuli, as well as an infrared LED and infrared phototransistor for detecting rats' interactions with the port. The port in the bottom row contained a stainless steel tube for delivering water rewards. Each training session lasted for 90 minutes.

#### **3.2.2.2 Behavior**

Trials began with both yellow and blue LED turning on in the center port. This cued the animal to poke its nose into the center port and hold it there for 1 s, after which the center lights were turned off and the choice ports became illuminated. We refer to this period as the 'soft fixation' period, as the animal was allowed to withdraw any time after the initial poke. From here, 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. The choice ports will be triggered as long as the animal performed a second poke into the center port. All animals exploited the soft fixation strategy, albeit to different degrees individually. They tended to withdraw after the initial poke but stayed close to the center port during the soft fixation period (FIGURE 3.9B).

During the soft fixation period a tone played from both speakers, indicating the

lottery magnitude for that trial. There were 6 distinct frequencies indicating different lottery magnitudes (2.5 kHz – 20 kHz, 75 dB), and all rats had a positive frequency-to-magnitude mapping. The frequency of each lottery was around one octave away from the adjacent tones, making distinguishing the different offers perceptually easy (Dent et al., 2018). For all animals except for subject 2160, the surebet port was on the left and the lottery port was on the right. 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 the center port within 300 seconds from trial start, or it did not make a choice 30 seconds after fixation. Violations were excluded from all analyses, except where they are specifically mentioned.

In some sessions, ‘free’ trials were interleaved with the choice and forced trials. Free trials were similar to choice trials except at the end of fixation both left and right port were illuminated with blue LEDs. The animal would receive a medium-

sized reward (2 times the surebet) regardless of which port it chose. The free trials were randomly interleaved with the choice and forced trials, and were introduced only after all the experiments presented in FIGURE 3.1-3.7 were completed.

### **3.2.2.3 Training pipeline**

Animal training took place in two distinct phases: the operant conditioning phase and the risky choice phase. Briefly, in the operant conditioning phase, rats became familiar with the training apparatus and learned to poke into the reward port when illuminated. Trials began with the illumination of the 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. Initially, there were more forced trials than choice trials to help them understand the task. Once the animals reliably differentiated between the low and high lottery choice trials, more choice trials were added. The intermediate frequencies were added one by one, contingent upon good behavior in the choice trials with existing frequencies. The lottery probability and 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, the lottery probability would be decreased.

### **3.2.3 Experimental manipulations**

#### **3.2.3.1 Surgery**

Surgical methods were similar to those described in Erlich et al. (2015). The rats were anesthetized with isoflurane and placed in a stereotaxic apparatus (RWD Life Science Co.,LTD, Shenzhen). The scalp was shaved, washed with ethanol and iodopovidone and incised. Then, the skull was cleaned of tissue and blood. The stereotax was used to mark the locations of craniotomies for the left and right FOF and PPC, relative to Bregma on the skull. Four craniotomies and durotomies were performed and the skull was coated with a thin layer of C&B Metabond (Parkell Inc., NY). Guide cannula along with the injector (RWD Life Science Co.,LTD, Shenzhen) was inserted 1.5 mm into the cortex measured from brain surface for each craniotomy. The guide cannulae were placed and secured to the skull one at a time with a small amount of Absolute Dentin (Parkell Inc., NY). The injector was removed from each guide once the guide was secured to the skull. After all four guide cannulae were in place, more Absolute Dentin was applied to cover the skull and further secure the guide cannulae. Vetbond (3M, U.S.) was applied to glue the surrounding tissue to Absolute Dentin. The animals were given 7 days to recover on free water before resuming training.

#### **3.2.3.2 Cannulae**

All 8 rats were implanted bilaterally in FOF (+2 AP,  $\pm 1.5$  ML mm from Bregma) with 26 AWG guide cannulae (RWD Life Science Co.,LTD, Shenzhen) and in lateral PPC (-3.8 AP,  $\pm 3.0$  ML from Bregma) with 26 AWG guide cannulae (4 cannulae per

rat total). The tip of the guide sat on the brain surface, while the 33 AWG injector was extended 1.5 mm below the bottom of the guide cannula. The dummy extended 0.5 mm into the cortex.

### **3.2.3.3 Infusions**

Infusions were performed once a week with normal training days taking place on all other days. This was to minimize adaptation to the effects of the muscimol and to have stable performance in the sessions immediately before infusion sessions. Animals were held by an experimenter during the infusion, no general anesthetic was administered. On an infusion day, the rat was placed on the experimenter's lap and the dummy cannulae were gently removed and cleaned with iodine and alcohol. The injector was inserted into the target guide cannula and reached 1.5 mm into cortex. A 1  $\mu$ L syringe (Gaoge, Shanghai) connected via tubing filled with mineral oil to the injector was used to infuse 0.3  $\mu$ L of muscimol (of various concentrations) into cortex. The injection was done over 1 minute, after which the injector was left in the brain for 5 more minutes to allow diffusion before removal. The thoroughly cleaned dummies were placed into the guide cannula. The rats began training 2 to 53 minutes after the infusion, the average time between infusion and starting of the behavioral session was 27 minutes. See FIGURE 3.8 for the complete list of all infusion doses, regions, and order for each rat.

### **3.2.4 Analysis**

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. Unless otherwise specified, the ‘control’ sessions refer to the sessions one day before any infusion event during the course of the experiment. All analysis and statistics were computed in R (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria), except for the biophysical model, which was simulated in Julia (1.6.0, Bezanson et al., 2017).

### 3.2.4.1 Generalized linear mixed-effects models (GLMM)

GLMMs were fit using the `lme4` R package (Bates et al., 2015). To test whether bilateral infusion had any effects on performance, we specified a mixed-effects model where the probability of a lottery choice was a logistic function of  $EV_{lottery} - EV_{surebet}$ , muscimol dosage ( $\mu\text{g}$ ) and their interaction as fixed effects. The rat and an interaction of rat,  $EV_{lottery} - EV_{surebet}$  and dosage were modelled as within-subject random effects. The expected value of lottery is the product of the lottery magnitude and lottery probability ( $EV_{lottery} = P_{lottery} \cdot V_{lottery}$ ). Similarly,  $EV_{surebet}$  denotes the expected value of surebet, which is simply the value of surebet here ( $EV_{surebet} = V_{surebet}$ , since  $P_{surebet} = 1$ ). In standard R formula syntax:

$$\text{chose\_lottery} \sim \text{delta\_EV} * \text{dosage} + (\text{delta\_EV} * \text{dosage} | \text{subjID}) \quad (3.1)$$

where `chose_lottery` is 1 if lottery was chosen on a trial, `delta_EV` is  $EV_{lottery} - EV_{surebet}$  and `subjID` is the subject ID for each rat.

To test whether unilateral infusions caused a left/right bias (as in Erlich et al., 2015), we specified a mixed-effects model similar to the one described above:

$$\text{choose\_right} \sim \text{rl\_delta\_EV} * \text{infusion\_side} + (\text{rl\_delta\_EV} * \text{infusion\_side} | \text{subjid}) \quad (3.2)$$

where `choose_right` is 1 if the right port is chosen on this trial, `rl_delta_EV` is  $EV_{right} - EV_{left}$  and `infusion_side` is a categorical variable with three levels: left, right and control. The plots in FIGURE 3.11 (FOF) and FIGURE 3.12 (PPC) show that the model fits for each rat are good, reflecting how the random effects allow for each rats' data to be fit, while also finding significant fixed effects.

To estimate the shift in indifference point induced by bilateral FOF inactivation, we first fit a GLMM as described above. We generated synthetic data points for `delta_EV` to extend its range, and the model was used to predict  $p(\text{Choose Lottery})$  for each synthetic data point. For each animal, we identified the `delta_EV` values that resulted in  $p(\text{Choose Lottery})$  to be between 0.499 and 0.501, which is the definition of indifference point. The average indifference point was obtained by taking the mean of such values across animals.

To test whether unilateral PPC infusions led to an ipsilateral bias in both free choice and risk choice trials, we specified a GLMM as following:

$$\text{choose\_ipsi} \sim \text{infusion} + (\text{infusion} | \text{subjid}) \quad (3.3)$$

where `choose_iphi` is a binary variable indicating whether the animal chose the side ipsilateral to the infusion side or not, and `infusion` is a binary variable representing the presence of an unilateral PPC infusion.

To estimate changes in reaction time, we used Linear Mixed-Effects Models (LMM). The formula for bilateral infusion is:

$$\log(RT) \sim \text{delta\_EV} * \text{dosage} * \text{choice} \quad (3.4)$$

$$+ (\text{delta\_EV} * \text{dosage} * \text{choice} | \text{subjID}) \quad (3.5)$$

where  $\log(RT)$  denotes the logarithm of reaction time, **choice** is a binary value for the surebet/lottery choice (0/1). Similarly, the formula for unilateral infusion is:

$$\log(RT) \sim \text{rl_delta_EV} * \text{infusion_side} * \text{choice} \quad (3.6)$$

$$+ (\text{rl_delta_EV} * \text{infusion_side} * \text{choice} | \text{subjID}) \quad (3.7)$$

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 mixed-effects model was specified:

$$\text{choose\_lottery} \sim \text{delta\_ev} + \text{prev\_outcome} + (\text{delta\_ev} + \text{prev\_outcome} | \text{subjID}) \quad (3.8)$$

where **prev\_outcome** is a categorical variable with three levels of previous outcome as above.

### 3.2.4.2 The three-agent mixture model

We developed a three-agent mixture model that used 4 parameters to transform the offers on each trial into a probability of choosing lottery as a weighted outcome of three agents (FIGURE 3.4A): a rational agent, a ‘lottery agent’ and a ‘surebet’ agent. For the rational agent, we assume an exponential term  $\rho$  for the utility function,  $U = V^\rho$ . A concave utility function ( $\rho < 1$ ) implies risk aversion, a linear function with  $\rho = 1$  implies being risk-neutral and a convex function ( $\rho > 1$ ) implies risk seeking. We modeled noise in the rational agent,  $\sigma$ , as the standard deviation in the expected utility distribution. This noise term represents the idea that animals’ internal representation of utility is noisy. The source of the noise could be neural noise or uncertainty in learning the map between perceptual stimuli and lottery magnitudes. Although perceptual noise could also play a role, we think the contribution would be small since the lottery tones are spaced around 1 octave apart from each other, which is easily distinguishable for rats (Dent et al., 2018). Concretely,

$$EU_L \sim \mathcal{N}(V_L^\rho P_L, \sigma) \quad (3.9)$$

$$U_{SB} \sim \mathcal{N}(V_{SB}^\rho, \sigma) \quad (3.10)$$

where the expected utility of lottery,  $EU_L$ , and the utility of the surebet,  $U_{SB}$  are Normal distributions.  $V_L, V_{SB}$  refer to the magnitude of lottery and surebet and  $P_L$  is the probability of lottery payout. The probability of choosing lottery for the rational agent then becomes

$$p_{\text{Choose Lottery}}^{\text{rational}} = p(EU_L > U_{SB}) \quad (3.11)$$

$$= p(EU_L - U_{SB} > 0) \quad (3.12)$$

$$= p(\mathcal{N}(V_L^\rho P_L, \sigma) - \mathcal{N}(V_{SB}^\rho, \sigma) > 0) \quad (3.13)$$

$$= p(\mathcal{N}(V_L^\rho P_L - V_{SB}^\rho, \sqrt{2}\sigma) > 0) \quad (3.14)$$

$$= 1 - \Phi(0; V_L^\rho P_L - V_{SB}^\rho, \sqrt{2}\sigma) \quad (3.15)$$

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. Note that this provides fits with similar likelihood as the softmax choice function with  $\beta$  as temperature:

$$\Delta EU = V_L^\rho P_L - V_{SB}^\rho \quad (3.16)$$

$$p(\text{Choose Lottery}) = \frac{1}{1 + e^{-\beta \Delta EU}} \quad (3.17)$$

The other two agents in the three-agent mixture model are the lottery and surebet agents. They represent the habitual bias of the animal to make one or the other choice regardless of the lottery offer, similar to biased lapse terms in Erlich et al. (2015). The probability of choosing lottery for the lottery agent is  $p_{\text{Choose Lottery}}^{\text{lottery}} = 1$  and for the surebet agent is  $p_{\text{Choose Lottery}}^{\text{surebet}} = 0$ .

The last step is to obtain  $p(\text{Choose Lottery})$  by mixing the probability from each agent  $\vec{P}$  with their respective mixing weights  $\vec{\omega}$  that sum up to 1. Formally,

$$P(\text{Choose Lottery}) = \vec{P} \cdot \vec{\omega} \quad (3.18)$$

$$= P_{\text{Choose Lottery}}^{\text{rational}} \omega_{\text{rational}} + 1 \cdot \omega_{\text{lottery}} + 0 \cdot \omega_{\text{surebet}} \quad (3.19)$$

$$\sum \vec{\omega} = 1 \quad (3.20)$$

### 3.2.4.3 Model 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. The prior over the utility exponent  $\rho$  was  $\text{Lognormal}(\log(0.9), 0.4)$ , a weakly informative prior that prefers  $\rho$  to be close to risk-neutral. The prior over noise  $\sigma$  was  $\text{Gamma}(6, 3)$ . The prior over the mixing weights  $\vec{\omega}$  was a Dirichlet distribution with the concentration parameter  $\alpha = [6, 2, 2]$ . The resulting  $\omega_{\text{rational}}$  distribution was broad and had the mean of 0.6, both  $\omega_{\text{lottery}}$  and  $\omega_{\text{surebet}}$  distribution had the mean of 0.2. By attributing more weight to the rational agent over the habitual agents, the prior reflected our selection of the experimental animals - only the ones with good psychometric curves were included. Four Markov chains with 1000 samples each were obtained for each model parameter after 1000 warm-up samples. The  $\hat{R}$  convergence diagnostic for each parameter was close to 1, indicating the chains mixed well.

### 3.2.4.4 Inactivation mixture model

We constructed a different version of the three-agent model, which considered two datasets from each rat simultaneously: an inactivation (bilateral FOF) and a control dataset. The model's raw parameters included  $\rho_{base}$ , a parameter for  $\rho$  in the log space, its prior was  $Lognormal(\log(0.9), 0.4)$ ;  $\sigma_{control}$ , as  $\sigma$  in the original version, with a prior of  $Gamma(6, 3)$ ;  $\omega_1$ , with a prior of  $\mathcal{N}(0, 2)$ , equivalent to  $\omega_{rational}$  after a *logistic* transformation; and  $\omega_2$  with a prior of  $\mathcal{N}(0, 2)$ , representing the proportion of the surebet agent in  $1 - logistic(\omega_1)$  after the *logistic* transformation itself, where  $logistic(x) = 1/(1 + e^{-x})$ :

$$\rho = e^{\rho_{base}} \quad (3.21)$$

$$\sigma = \sigma_{control} \quad (3.22)$$

$$\omega_{rational} = logistic(\omega_1) \quad (3.23)$$

$$\omega_{surebet} = (1 - \omega_{rational}) \cdot logistic(\omega_2) \quad (3.24)$$

$$\omega_{lottery} = (1 - \omega_{rational}) \cdot (1 - logistic(\omega_2)) \quad (3.25)$$

For the inactivation dataset, we added a new parameter for each raw parameter in order to estimate the effects of inactivation:

$$\rho = e^{\rho_{base} + \Delta\rho_{base}} \quad (3.26)$$

$$\sigma = \sigma_{control} \cdot \sigma_\pi \quad (3.27)$$

$$\omega_{rational} = logistic(\omega_1 + \Delta\omega_1) \quad (3.28)$$

$$\omega_{surebet} = (1 - \omega_{rational}) \cdot logistic(\omega_2 + \Delta\omega_2) \quad (3.29)$$

$$\omega_{lottery} = (1 - \omega_{rational}) \cdot (1 - logistic(\omega_2 + \Delta\omega_2)) \quad (3.30)$$

where  $\Delta\rho_{base}$  denotes the change in  $\rho$  in the log space, it had a prior of  $\mathcal{N}(0, 1)$ ;  $\sigma_\pi$ , with a prior of  $Lognormal(0, 0.1)$ , represents how the infusions could scale noise;  $\Delta\omega_1$  and  $\Delta\omega_2$  fit potential changes in  $\omega_1$  and  $\omega_2$  before the *logistic* transformation, respectively.

We constructed two other variants of the inactivation model for model comparison. For the  $\rho$ -only model, both  $\Delta\omega_1$  and  $\Delta\omega_2$  were fixed to be 0 during fitting. For the  $\omega$ -only model,  $\Delta\rho_{base}$  was fixed to be 0.

### 3.2.4.5 Synthetic datasets

To test the validity of our model, we created synthetic datasets with parameters generated from the prior distributions described above. The three-agent model was fit to the synthetic datasets, and it was able to recover the generative parameters accurately (FIGURE 3.15A). This assured that our model can capture the behavior well and has no systematic bias in estimating the parameters.

### 3.2.4.6 Mixture model prediction confidence intervals

To generate model predictions in between the actual lottery lottery magnitudes (as in FIGURE 3.4B), we generated a synthetic dataset with narrowly-spaced lottery magnitudes (incremented by 1). Then, we sampled parameters from the estimated posteriors and computed the probability of choosing the lottery given the synthetic offers. The resulting output is a `n_iter` × `n_lott_mag` matrix, where `n_iter` is the number of Markov samples and `n_lott_mag` is the length of unique lottery magnitudes. Finally, binomial 80%, 90%, and 95% confidence intervals for each lottery offer were estimated by taking the 10% and 90%, 5% and 95%, and 2.5% and 97.5% percentile of `n_iter` predicted choices, respectively.

### 3.2.4.7 Mixture model comparison

To understand which model describes the inactivation results best, we performed 10-fold cross-validation of the MCMC fits of each model. For each fold, the model first fit on the training data, containing 90% of the original data from each condition (control and bilateral FOF). We then computed (in the `generated quantities` block) the log predictive densities by passing in the held-out data, using the posterior 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

$$\widehat{lpd} = \sum_{i=1}^n \log \left( \frac{1}{S} \sum_{s=1}^S p(y_i | \theta^s) \right), \quad (3.31)$$

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

#### 3.2.4.8 Biophysical model

We generated a 6-node rate model to understand how muscimol inactivation of the FOF could cause a reduction in lottery choices via a change in the curvature of the utility function. The activity of the six nodes,  $X$ , are governed by the following equations, where  $v$  is the magnitude of the lottery and the  $i$  in  $g(v, t, i)$  represents the node index (1-6). Simulation was done using Euler's method in Julia (1.6.0, Bezanson et al., 2017):

$$dX = dt(-X/\tau + WX + g(v, t, i) + \mathcal{N}(0, \sigma)) \quad (3.32)$$

$$X = f(X + dX) \quad (3.33)$$

$$f(x) = \begin{cases} x & \text{if } 0 \leq x \leq 100 \\ 0 & \text{if } x < 0 \\ 100 & \text{if } x > 100 \end{cases} \quad (3.34)$$

$$g(v, t, i) = \begin{cases} 0 & \text{if } 0.1 > t > 1 \text{ s or } i \leq 3 \\ v & \text{otherwise} \end{cases} \quad (3.35)$$

$$\tau = 0.15 \text{ s} \quad (3.36)$$

$$dt = 0.001 \text{ s} \quad (3.37)$$

$$\sigma = \frac{5}{6} \cdot dt^{-1} \quad (3.38)$$

$$\text{for } w_{ij} \in W, \quad (3.39)$$

$$w_{ij} \sim \mathcal{N}(5/6, 0) \quad (3.40)$$

We began the simulation of each trial a few seconds before the input was turned on, to allow the network to reach its baseline fixed-point. We examined different instantiations of this model by generating the weight matrix,  $W$ , from different random seeds. Many (but not all) of these networks gave qualitatively similar results.

### 3.2.5 Surebet learning

To test the role of PPC in learning, we periodically changed the surebet magnitude in a model-based way to shift the decision boundary. For each shift, we fit the three-agent model on control data from the past 14 days to obtain a set of parameters. Using a binary search algorithm, we then used those parameters to generate synthetic choices with different surebet magnitudes until we found a value that produces a shift in overall probability choosing lottery ( $p(\text{Choose Lottery})$ ) close to the target (drawn uniformly from  $\pm U(0.2, 0.3)$ ). The new surebet magnitude was assigned to the animal on the day of change. All animals in the surebet learning experiment had undergone two rounds of shift without any infusion, in the course of 14 days, to acclimate them to the new routine before bilateral PPC infusions. The first two surebet change sessions are not included in the analysis of FIGURE 3.7.

## 3.3 Results

### 3.3.1 Behavior

In this paper, we only present behavior from sessions after the animal was implanted with cannulae for experiments. Unless otherwise specified, control sessions were the sessions from the day before the infusion sessions. The animals' choices were consistent with a utility-maximizing strategy: they had relatively few violations of first-order stochastic dominance (i.e. they chose the surebet when the lottery magnitude was less than the surebet magnitude) and they increased the proportion of lottery choices monotonically with increasing expected value (FIGURE 3.1B-D). Six

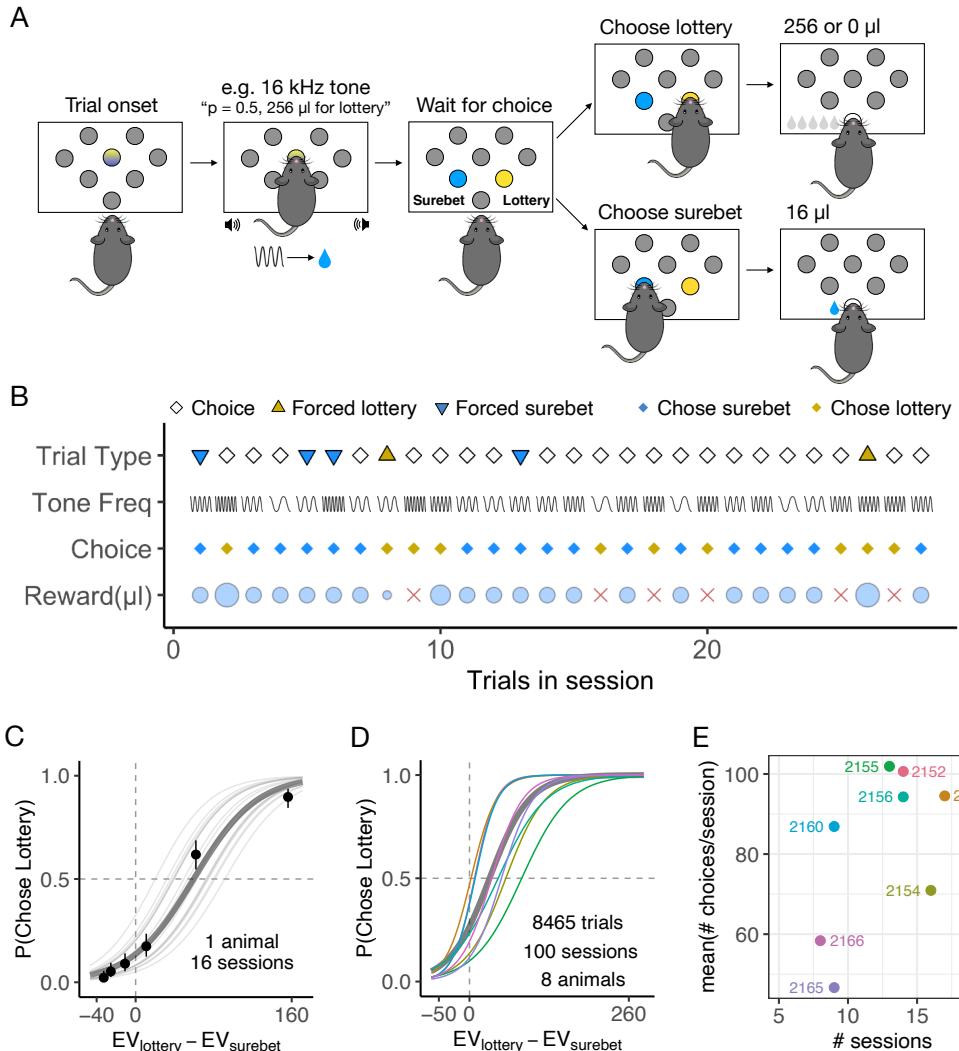
of the rats were risk-averse and two were close to risk-neutral (FIGURE 3.1D). On average, each rat completed 82 choices in a control session (FIGURE 3.1E).

### 3.3.2 Effects of silencing FOF and PPC on the risky choice task

All animals experienced three different types of inactivations (left, right and bilateral) in two brain areas (FOF & PPC). In total, we include 7456 choice trials from 127 infusions sessions into the FOF and PPC of 8 rats. The details of region, order and dosage of the infusions for each rat is shown in FIGURE 3.8.

#### 3.3.2.1 FOF silencing shifted choices away from the lottery

Bilateral silencing of the FOF (FIGURE 3.2A) resulted in a dose-dependent reduction in lottery choices (FIGURE 3.2B; for individual subjects see FIGURE 3.11A). A generalized-linear mixed-effects model (GLMM) of the bilateral infusions found a significant main effect of muscimol dosage ( $\beta_{dose} = -3.18 \pm 0.92, p < 0.01$ ). The mean indifference point (in units of  $EV_{lottery} - EV_{surebet} = \mu L$  of water) shifted from  $50.92 \pm 11.56$  in control to  $154.43 \pm 23.49$  under  $0.3 \mu g$  muscimol ( $T_8 = -3.95, p < 0.001$ ). In other words, inactivating bilateral FOF is equivalent to adding around  $100 \mu L$  to the surebet. There also was a small but significant decrease in the slope of the logistic curve ( $\beta_{EV_{lottery}-EV_{surebet}:dose} = -0.08 \pm 0.02, p < 0.001$ ). Bilateral silencing of the FOF did not consistently change animal's reaction time, defined as the time from center port withdrawl until a choice port poke (Linear mixed-effects model, LMM,  $\beta_{dose} = 0.21 \pm 0.29, p = 0.574$ ). However, there was a significant slowing effect in three animals: 2152 ( $\beta_{dose} = 2.33 \pm 0.42, p < 0.001$ ), 2153 ( $\beta_{dose} = 1.31 \pm 0.39, p < 0.01$ )



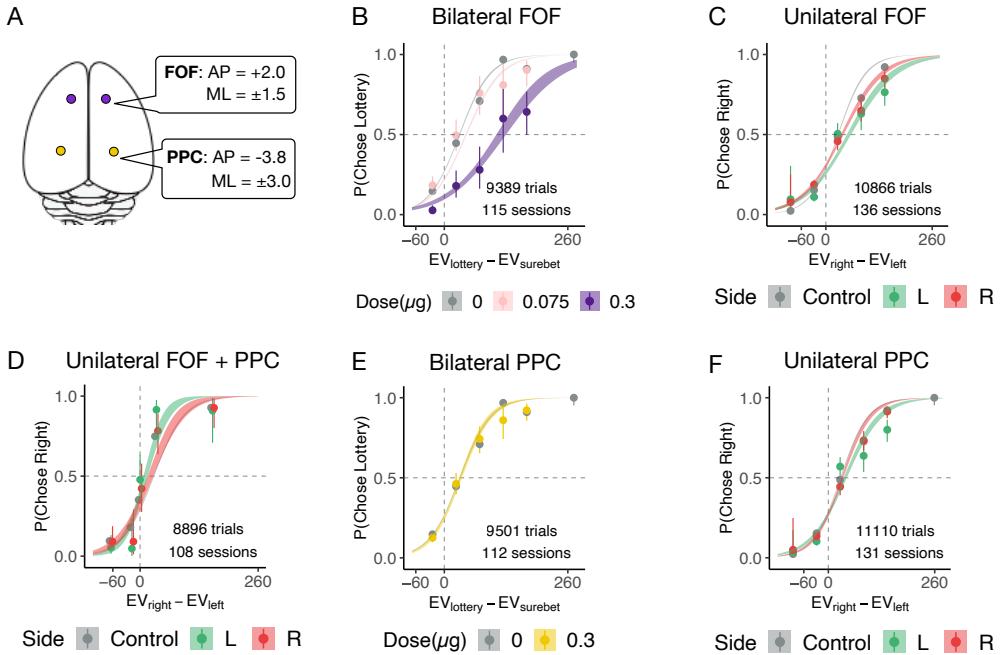
**Figure 3.1.** The risky choice task and animal behavior. **A.** Schematic of the risky choice task. **B.** Timeline of trials in one example session as in FIGURE 2.1B. **C.** Example subject performance from 16 control sessions 1 day before an infusion. The probability of choosing lottery is plotted as a function of the expected value of lottery minus the expected value of surebet ( $V_{lottery} - V_{surebet}$ ), where  $V$  represents  $\mu\text{L}$  of water. The circles with error bars are the mean and 95% binomial confidence intervals. The lines are generated by a generalized linear model, the thin gray lines are fit to each session, the thick gray line fit to all the sessions combined. **D.** Subject performance from 100 control sessions 1 day before an infusion event (8 rats). The colored lines are generated by a generalized linear model fit to each animal, the thick gray line fit to all the sessions combined. **E.** The number of control sessions within the infusion period, and the average number of choice trials, colored by subject.

and 2166 ( $\beta_{dose} = 1.22 \pm 0.25, p < 0.001$ ), possibly due to muscimol spillover into the adjacent M1 area (FIGURE 3.10C). Overall, the slowing effect from bilateral FOF inactivation was less reliable across animals than the effect on choice (FIGURE 3.13A), suggesting the effect on choice was not primarily driven by changes in movement.

Unilateral infusions had a smaller effect compared to bilateral infusions (FIGURE 3.2C). Infusions of  $0.3 \mu\text{g}$  muscimol into the left and right FOF resulted in small but significant decrease in the slope of the *logistic* curve ( $\beta_{EV_{right}-EV_{left}:left} = -0.01 \pm 0.003, p < 0.001$ ;  $\beta_{EV_{right}-EV_{left:right}} = -0.01 \pm 0.004, p = 0.049$ ). These results were surprising for two reasons. First, we expected an ipsilateral bias, but both left and right infusions shifted animals slightly to choose leftward choices. As seven out of eight animals had the surebet port assigned on the left, it is possible that the decrease in choosing *right* after silencing either side of the FOF was, in fact, a partial effect of bilateral FOF inactivation (decrease in choosing the lottery). Second, these effects are very weak compared to the large ipsilateral biases caused by unilateral FOF silencing in previous tasks (Erlich et al., 2015, Kopec et al., 2015, Erlich et al., 2011). The discrepancy may be due to the memory component in their tasks, whereas our task does not have one. Overall, unilateral infusions in FOF did not change animal's reaction time (LMM,  $\beta_L = 0.07 \pm 0.07, p = 0.44$ ;  $\beta_R = 0.01 \pm 0.07, p = 0.89$ ).

### 3.3.2.2 PPC silencing had minimal effect on the risky choices

Bilateral silencing of PPC resulted in minimal effect on the risky-choice behavior ( $\beta_{dose} = -0.60 \pm 0.57, p = 0.29$ ; FIGURE 3.2E). The lack of main effect of dose was found in 5 out of 7 animals (FIGURE 3.12A), except for 2153 ( $\beta_{dose} = -2.12 \pm 0.59, p < 0.001$ ) and 2156 ( $\beta_{dose} = -2.19 \pm 0.84, p < 0.01$ ). To test for any lateralized



**Figure 3.2.** Bilateral and unilateral infusions in FOF and PPC. **A.** Top-down view of the rat cortex with the target coordinates of FOF and PPC, where the cannulae were implanted. **B.** Bilateral infusion of muscimol (0.3  $\mu$ g) into the FOF significantly shifted the choices towards the lottery. Control sessions 1 day before an infusion are shown in gray ( $n = 17$  sessions, 8 rats), 0.075  $\mu$ g per side bilateral FOF infusions ( $n = 6$  sessions, 5 rats) are in light pink, 0.3  $\mu$ g per side bilateral FOF infusions ( $n = 9$  sessions, 8 rats) are in dark purple. The circles with error bars are the mean and 95% binomial confidence intervals. The ribbons are from a generalized linear model fit to the data. See details in Methods. **C.** Unilateral infusion of muscimol into the left and right FOF resulted in a small but reliable shift towards surebet. Control sessions are in gray ( $n = 39$  sessions, 8 rats), 0.3  $\mu$ g left FOF infusions ( $n = 16$  sessions, 8 rats) are in green, 0.3  $\mu$ g right FOF infusions ( $n = 20$  sessions, 8 rats) are in red. **D.** Simultaneous unilateral inactivation of FOF and PPC had no effect. Control sessions are in gray ( $n = 8$  sessions, 4 rats), 0.3  $\mu$ g left FOF infusion and 0.6  $\mu$ g left PPC infusions ( $n = 4$  sessions) are in green, 0.3  $\mu$ g right FOF infusion and 0.6  $\mu$ g right PPC infusion ( $n = 4$  sessions) are in red. **E.** Bilateral infusion of muscimol into the PPC had no effect. Control sessions are in gray ( $n = 24$  sessions, 7 rats), 0.3  $\mu$ g per side bilateral PPC infusions ( $n = 12$  sessions, 7 rats) are in gold. **F.** Unilateral infusion of muscimol into the PPC had no effect. Control sessions are in gray ( $n = 31$  sessions, 8 rats), 0.3  $\mu$ g left PPC infusions ( $n = 11$  sessions, 7 rats) are in green, 0.3  $\mu$ g right PPC infusions ( $n = 19$  sessions, 8 rats) are in red.

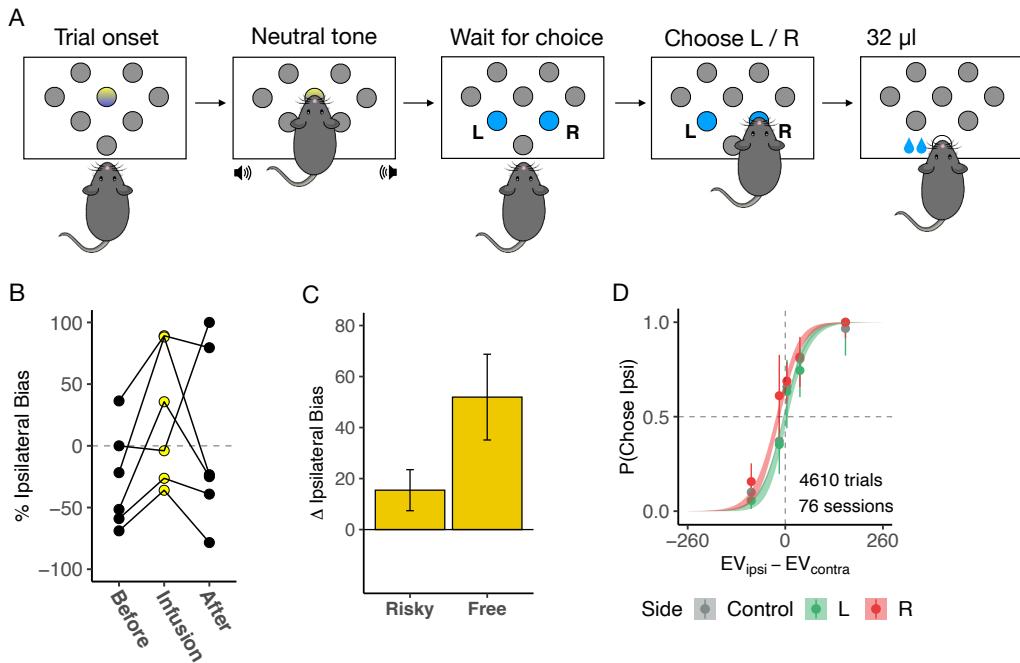
effects from unilateral PPC infusions, we performed a second GLMM test where the choice on each trial was a logistic function of  $EV_{right} - EV_{left}$ , infusion side and their interaction as fixed effects. No significant effects were found on the group level (FIGURE 3.2F, all  $p > 0.5$ ). This was consistent across all subjects except for 2152, 2153 and 2155, left infusion produced a small but significant shift away from lottery for these animals (FIGURE 3.12B). To probe whether perturbation of FOF could reveal an effect of PPC inactivation, we inactivated unilateral FOF ( $0.3 \mu\text{g}$ ) while unilaterally inactivating PPC with  $0.6 \mu\text{g}$  muscimol. The simultaneous inactivation, still, had no significant effect on the behavior (FIGURE 3.2D, all  $p > 0.5$ ). Overall, the results suggest that PPC inactivation was ineffective in biasing the risky choices. Thus, our hypothesis that the PPC may be involved in economic decisions because they are an expression of an internal preference was not supported.

In order to establish that our infusions into PPC were effective, after completing all of the experiments reported related to risky-choice, we added a ‘free’ trial type (as in Erlich et al., 2015). On a free trial, both the surebet port and lottery port were illuminated with blue-LEDs after fixation, accompanied by a brief neutral tone. The animals were rewarded twice the magnitude of the surebet reward regardless of which port they chose (FIGURE 3.3A). These types of trials have been demonstrated to be sensitive to unilateral silencing of the PPC (Erlich et al., 2015, Katz et al., 2016). We randomly intermixed 11% free trials with 22% forced trials and 67% choice trials on the control days. After a few sessions with the new trial type, rats expressed a consistent bias on the free trials and still performed the choice trials in a utility-maximizing way. The proportion of free trials was increased to 50% on the infusion day, with the rest being 12.5% forced trials and 37.5% choice trials. Infusions

of muscimol ( $0.6 \mu\text{g}$ ) into one hemifield of PPC (opposite to the animal’s preferred side) produced a substantial ipsilateral bias on free trials (FIGURE 3.3B;  $\beta_{infusion} = 1.19 \pm 0.50, p < 0.05$ ). The ipsilateral bias in free trials was observed even while, consistent with our previous PPC inactivation results, there was no ipsilateral bias on the interleaved choice trials (FIGURE 3.3D;  $\beta_{infusion} = 0.18 \pm 0.14, p = 0.189$ ). These free trial inactivation results provide a clear positive control for our PPC inactivations, demonstrating that the lack of effect on choice trials was not caused by a technical issue (like clogged cannula).

### 3.3.3 A three-agent mixture model of risky choice

While the effects of FOF silencing confirmed its role in decisions under risk (FIGURE 3.2B), the GLMM results did not provide insight into the specific role that the FOF might play. To better understand animal behavior in the task and the role of the FOF, we developed a three-agent mixture model (FIGURE 3.4A). The first agent is a ‘rational’, utility-maximizing agent (Von Neumann and Morgenstern, 1953) with two parameters:  $\rho$ , which controls the shape of the utility function ( $U = V^\rho$ );  $\sigma$ , which captures the decision noise. The other two agents were stimulus-independent agents which either habitually chose the lottery or the surebet. The relative influence of the agents is controlled by their mixing weights  $\omega$ , 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. We estimated the joint posterior over the parameters for each subject separately using Hamiltonian Monte Carlo sampling in Stan (Carpenter et al., 2016) and validated that the model can correctly recover generative parameters from synthetic data (FIGURE 3.15A). Details of the modeling, including the priors, can



**Figure 3.3.** Unilateral PPC inactivation induced an ipsilateral bias on free trials. **A.** Schematic of the free trials. After fixation at the center port accompanied by a neutral tone, the animal was free to choose the left or right port, both illuminated in blue LEDs. Choosing either port resulted in a reward twice the magnitude of surebet. The free trials were randomly interleaved with the forced and choice trials. **B.** Unilateral PPC infusions ( $0.6 \mu\text{g}$ ) led to a significant ipsilateral bias towards the side of infusion. This panel shows % ipsilateral bias:  $(\sum \text{choose\_infusion\_side} - \sum \text{choose\_other\_side}) / \sum \text{total\_choices}$ , when the side of infusions was chosen to be the opposite to the animals' preferred side. % ipsilateral bias was computed using free trials from the previous 3 sessions, the infusion session, and the following 3 sessions for 6 subjects. **C.** Unilateral PPC infusions generated a significant  $52 \pm 16\%$  (mean  $\pm$  s.e. across rats,  $n = 6$ ) change in % ipsilateral bias on free trials compared to control sessions (3 pre-infusion sessions). For the choice trials from the same sessions, the change in % ipsilateral bias was not significant ( $15 \pm 8\%$ ). **D.** Performance on the choice trials was not affected. Control sessions from the 3 pre-infusion sessions ( $n = 65$  sessions, 6 rats) are in gray,  $0.6 \mu\text{g}$  left PPC infusions ( $n = 5$  sessions) are in green,  $0.6 \mu\text{g}$  right PPC infusions ( $n = 6$  sessions) are in red.

be found in the Methods section. The motivation for developing the mixture model was that the animals' choices, while clearly sensitive to the lottery offer, showed some stimulus-independent biases. In other words, even for the best lottery they sometimes chose the surebet and for the worst lottery (which had a value of 0) they sometimes chose the lottery. For example, subject 2156 has a psychometric curve that asymptotes in a way that is inconsistent with a pure utility-maximizing strategy (FIGURE 3.4B).

Trial history effects could have been incorporated by allowing model parameters to vary depending on the outcome of the previous trial (as in Constantinople et al., 2019b). However, our animals seemed to understand that the lottery offer was independent across trials, and we did not see any significant effects of previous trial's outcome on choice in control sessions ( $\beta_{lottery-win} = 0.20 \pm 0.12, p = 0.08$ ;  $\beta_{lottery-lose} = 0.17 \pm 0.09, p = 0.08$ ). For this reason, we decided to formulate the three-agent mixture model without trial history parameters. Our animals' behavior stands in contrast to a substantial number of published results demonstrating strong trial history effects in rodent decision-making even when the optimal strategy is to only use information on the current trial (e.g. Morcos and Harvey, 2016, Scott et al., 2015, Constantinople et al., 2019b). We speculate that an important difference is that in traditional rodent two-alternative forced-choice tasks, the rewards were delivered at the choice ports, but in our task all rewards were delivered at a single reward port (but for counter examples where there is history dependence despite using a single reward port, see Aguillon-Rodriguez et al., 2021, Zalocusky et al., 2016).

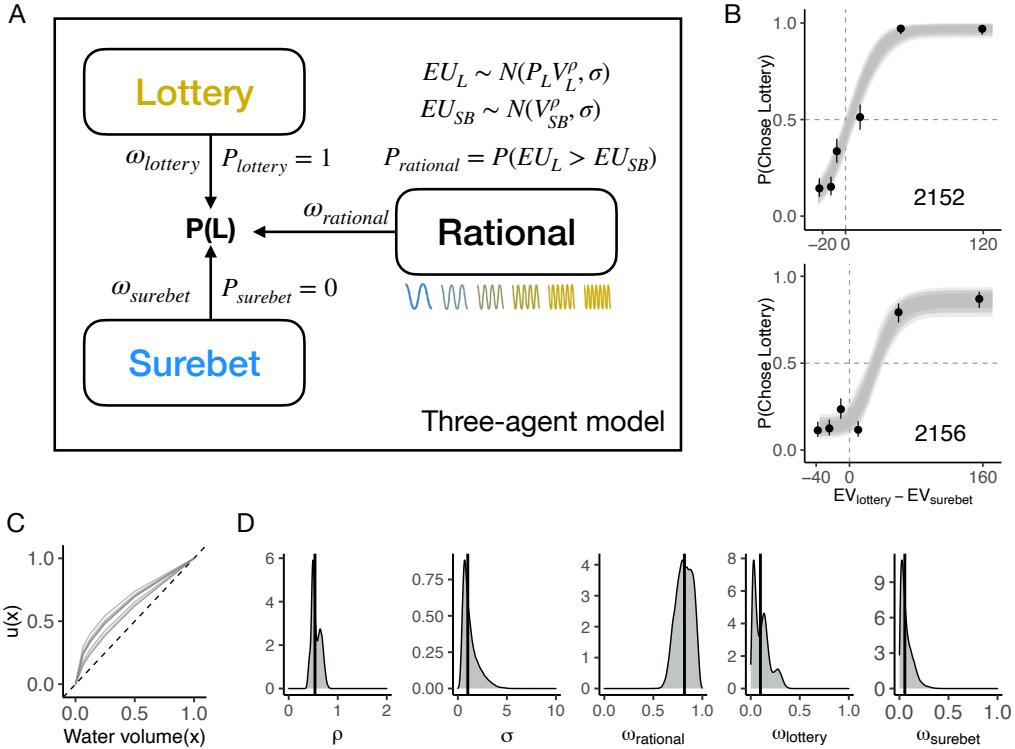
The three-agent model fit the control behavior well (see two example animals in FIGURE 3.4B, all animals in FIGURE 3.15B). For example, rat 2156 chooses

the lottery about 10% of the time for the 4 worst lotteries. This behavior is not well described by a ‘pure’ utility-maximizing model. All animals had a decelerating utility function (95% C.I. of  $\rho < 1$  for all animals; FIGURE 3.4C). Note that the *effective* risk-preference is influenced by both  $\rho$  and  $\omega$ . For example, the indifference point of 2152 is close to 0, implying that it is effectively risk-neutral (FIGURE 3.4B). However, this comes from its bias towards choosing the lottery ( $\omega_{lottery} = 0.16$ ) balancing its decelerating utility function ( $\rho = 0.62$ ; Table 3.1). The animals had small but varying levels of decision noise ( $\sigma = 1.00 [0.35 3.40]$ , median and 95% C.I. of concatenated posteriors across animals), indicating that they were sensitive to water rewards just a few  $\mu\text{L}$  apart. Their choices were guided mostly by the rational agent ( $\omega_{rational} = 0.82 [0.65 0.95]$ ), with little influence from the lottery agent ( $\omega_{lottery} = 0.10 [0.01 0.31]$ ) and the surebet agent ( $\omega_{surebet} = 0.05 [0.01 0.23]$ ).

### 3.3.4 Bilateral FOF inactivation reduced the utility exponent

$$\rho$$

In order to quantify how the infusions influenced model parameters, we constructed a new version of the three-agent model that fit the 0.3  $\mu\text{g}$  bilateral FOF and 0.3  $\mu\text{g}$  bilateral PPC infusion data as perturbations of the control dataset for each subject (FIGURE 3.5A & FIGURE 3.18). We chose priors for the effects of perturbation such that the model favored no effect of inactivation (i.e. zero mean for shifts and one mean for scaling effects). Bilateral PPC infusion led to no reliable changes across subjects for all parameters, which was consistent with the results from the GLMM (FIGURE 3.18 & TABLE 3.1). From the GLMM (and visual inspection), we knew that bilateral FOF silencing substantially shifted the subjects to being ef-



**Figure 3.4.** The three-agent mixture model and model fits. **A.** The three-agent mixture model. The animal’s choice is modelled as a weighted average of the three agents, each implementing a different behavioral strategy to perform the task. Each agent outputs a probability of choosing lottery that makes up the probability vector  $\vec{P}$ , which is combined using their respective weights  $\vec{\omega}$ . See Methods for model details. **B.** The three-agent mixture model can fit the control behavior well. The circles with error bars are the binned mean and 95% binomial confidence intervals. The ribbons are model predictions generated using the fitted parameters. The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively. **C.** The subjective utility functions for each rat computed using maximum a posteriori  $\rho$  estimation, normalized by the maximum water volume. **D.** Density plots of concatenated posterior samples (4000 each) from the model fits across 8 animals. The black bar is the median of the distribution.



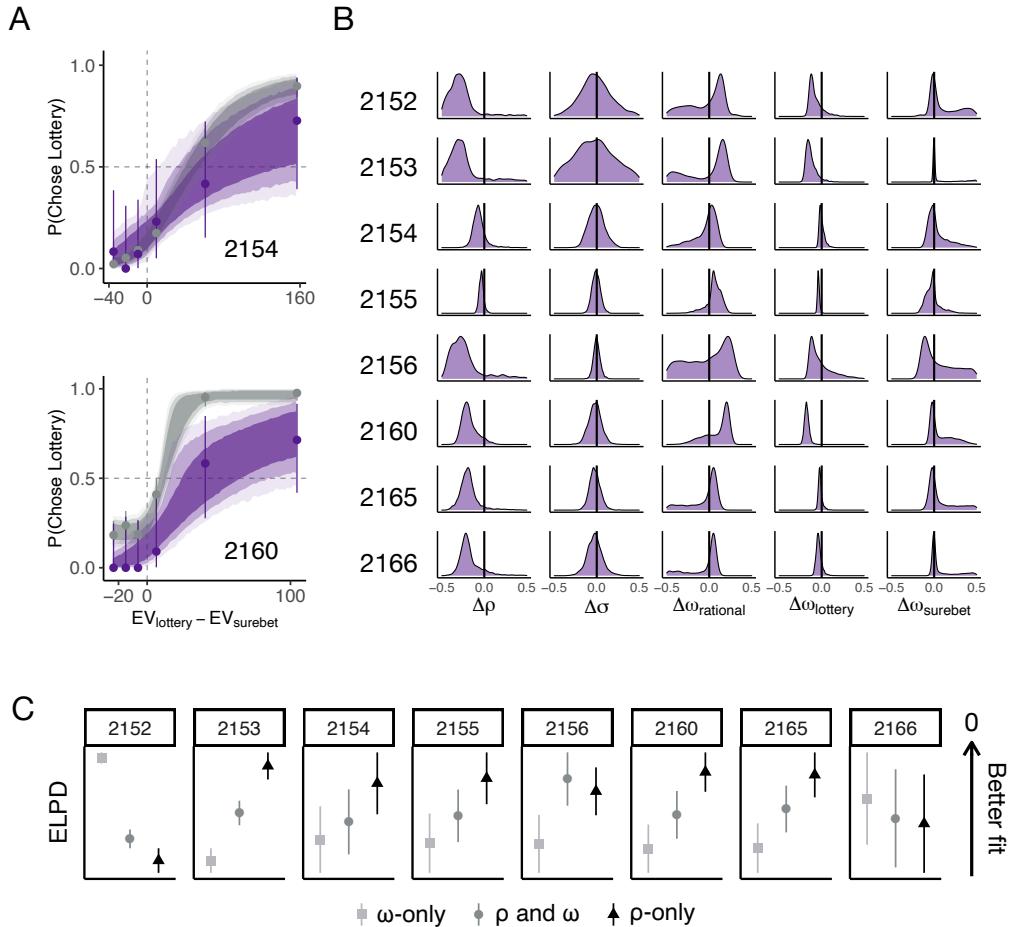
to change under inactivation. Similarly, the ‘ $\omega$ -only’ model had parameters allowing only  $\omega$  and  $\sigma$  but not  $\rho$  to change. The ‘ $\rho$  and  $\omega$ ’ model was the standard inactivation model that allowed every parameter to change under inactivation. We found that in 6 out of 8 subjects, model comparison result preferred the  $\rho$ -only model over the  $\omega$ -only model (FIGURE 3.5C). The  $\omega$ -only model was strongly preferred only in one subject’s dataset (2152). Taken together, these results suggest that the most parsimonious interpretation of the inactivation-induced effect is a reduction in the utility exponent,  $\rho$ .

How can silencing the FOF change the exponent of the utility function? Previous silencing and modeling results suggested that the FOF is part (1/6) of a distributed circuit for maintaining a prospective memory of choice (Kopec et al., 2015). Inspired by that finding, we constructed a 6-node rate model of a distributed circuit for encoding action-value, where the FOF represented one node in that network (FIGURE 3.6A Burak and Fiete, 2012). Three nodes other than the FOF node received input representing the magnitude of the lottery. The all-to-all weight matrix was generated randomly, but the distribution of the weights was chosen such that the response of the network to the inputs was in the dynamic regime of the nodes ( $0 < Hz < 100$ ). Other network parameters (noise  $\sigma$  and time-constant  $\tau$ ) were chosen to generate a control network response with reasonable dynamics (FIGURE 3.6B) that encoded the lottery value in the population activity of the network (FIGURE 3.6C, gray circles). In this regime, we found that silencing the FOF node scaled down the network’s responses. We can think of this network as encoding the expected utility of choosing the lottery by transforming the lottery sound into *utils* (encoded as spike rate). At the time of the go-cue, this activity could become bistable: where the utility of the

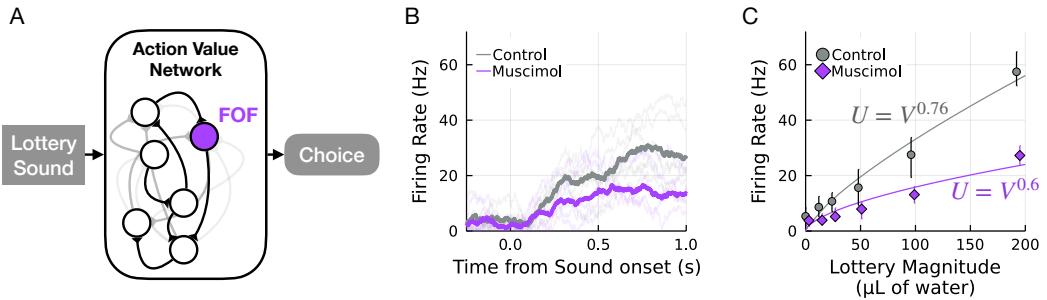
surebet determines the unstable fixed point (similar to Machens, 2005). Alternatively, a downstream region could compare the output of this network with the remembered surebet utility. In any case, scaling down the input-output transform of the network (FIGURE 3.6C, purple circles) would shift the indifference point (the lottery that had the same activity level as the surebet comparator), which would, behaviorally, appear as a change in the power-law utility function  $U = V^\rho$ . For the control network, the network approximates a function with  $\rho \approx 0.76$ . After silencing the FOF node, the exponent of the utility functions shifted down,  $\rho \approx 0.6$  (FIGURE 3.6C). This biophysical model provides a mechanistic explanation for our finding that silencing the FOF with muscimol caused animals to avoid choosing the lottery (FIGURE 3.2B) through a change in the exponent of the utility function (FIGURE 3.5B).

### 3.3.5 Bilateral PPC inactivation did not impair learning

The GLMM analysis above shows that the rat PPC was not causally involved in the risky choice task. However, numerous studies have found that neural activity in PPC correlated with decision variables in both perceptual and economic tasks. The question thus remains, what is the purpose of these decision-related signals in PPC? Recently, Zhong et al. (2019) found that PPC silencing impaired the ability of mice to re-categorize previously experienced stimuli based on a new category boundary in an auditory decision-making task. Moreover, after the stimuli were re-categorized, PPC activity was no longer required for performance. Motivated by their findings, we tested whether PPC was necessary for re-categorizing stimuli in our task. To do so, we employed a model-based change in the surebet magnitude that effectively shifted the decision boundary without changing the frequency-to-lottery mapping (FIGURE



**Figure 3.5.** Bilateral FOF inactivation reduced the utility exponent. **A.** Psychometric curves for two example animals. The circles with error bars are the binned mean and 95% binomial confidence intervals. The ribbons are model predictions generated using the fitted parameters (control—gray, bilateral FOF inactivation—purple). The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively. All animals are shown in FIGURE 3.17. **B.** Normalized (with peak at 1) density plots showing deviation of the model parameters under bilateral FOF inactivation from control. The most reliable finding across animals is that  $\Delta\rho$  is negative. **C.** Ten-fold cross-validation results comparing three model variants: where  $\omega$  but not  $\rho$  was allowed to change under the inactivation dataset, both  $\rho$  and  $\omega$  were allowed to change, and  $\rho$  but not  $\omega$  was allowed to change. The points with error bars are the expected log posterior density (ELPD) and its standard error on each animal’s dataset. The  $\rho$ -only model was preferred to the  $\omega$ -only model in 6 out of 8 animals.



**Figure 3.6.** Biophysical model of FOF silencing. **A.** We implemented a 6-node rate model of a distributed action-value network with random connectivity ( $W_{ij} \sim \mathcal{N}(5/6, 1)$ ). The FOF is 1 of the 6 nodes (in purple). The input to the network was the lottery magnitude. For the following plots the random seed for generating  $W$  was set to 131 and then that  $W$  was used for all further simulation, but similar results can be obtained with other  $W$  generated with the same statistics from a different seed. **B.** Example of the network response to lottery sound with magnitude of 96  $\mu\text{L}$  under control conditions (with all the nodes active, in grey) and under FOF silencing (the FOF node is set to zero, in purple). The dark traces represent the mean network activity and the light traces represent the activity of the 6 individual nodes. **C.** Silencing FOF scales down the representation of the action-value of the lottery, which could explain the shift in  $\rho$ . We ran the network for 20 ‘trials’ of each lottery  $\in [0, 12, 24, 48, 96, 192] \mu\text{L}$ . The grey circle are the mean and 95% CI for the network response in the control conditions and the purple diamonds are the mean and 95% CI for the network response when the FOF node is silenced. Fitting a power-law utility function,  $U = V^\rho$  to the network activity gives  $\rho \approx 0.76$  for control, and after FOF silencing  $\rho \approx 0.6$ . The thin lines are power-law utility functions that approximate the transformation from units of reward ( $\mu\text{L}$ ) to *utils* in spikes / second.

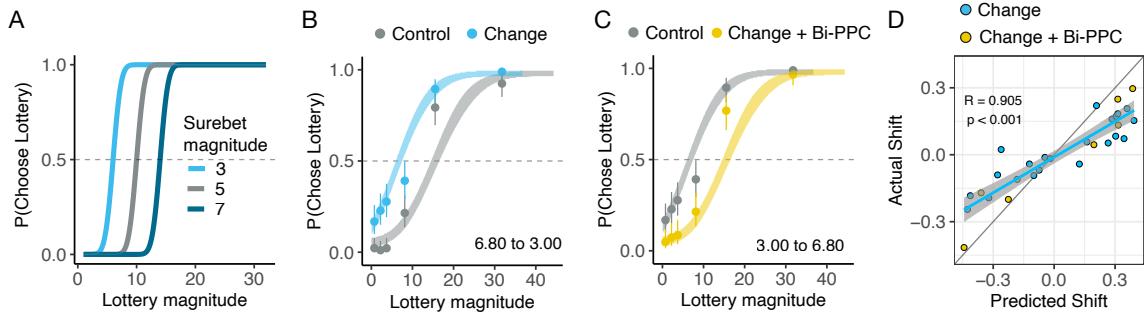
3.7A). As such, some frequencies that were previously preferred over the surebet became unpreferred (and vice-versa, depending on the direction of the shift). To estimate the required shifts, we first fit the three-agent model on data from the past 14 sessions. We then used the fit to generate synthetic choices on different surebet magnitudes, until we found the one that resulted in a shift in the overall probability of choosing lottery ( $p(\text{Choose Lottery})$ ) close to the target (drawn uniformly from  $\pm U(0.2, 0.3)$ ; see details in Methods). To familiarize animals with the new paradigm, their surebet magnitudes were changed weekly for two weeks prior to any infusion. Two out of six animals failed to show appropriate adaptation of behavior following change in surebet magnitude; they were excluded from analysis in this section. The other four animals reliably shifted their choices more towards surebet when its magnitude increased, and more towards lottery when its magnitude decreased (See example animal in FIGURE 3.7B, all other animals in FIGURE 3.19)

After two weeks, on the day of surebet change, we infused 0.6  $\mu\text{g}$  muscimol into each side of PPC in these four animals. The animals learned the new surebet magnitude and adjusted their behavior to the same extent in control and PPC inactivation sessions (see example animal in FIGURE 3.7C). To validate that the animals adapted in a preference-preserving manner, we fit one model to all the sessions with different surebet magnitudes for each animal. We then used these parameters to predict the shift in  $p(\text{Choose Lottery})$  on the day of change, and compared it against the actual shift in  $p(\text{Choose Lottery})$  in that session (FIGURE 3.7D). The model was able to accurately predict the shift ( $R = 0.905, p < 0.001$ ), suggesting that the animals adapted their choices without altering the underlying preference and strategy. Bilateral PPC inactivation did not impair the learning of new surebet magnitudes, as the shifts in

$p(\text{Choose Lottery})$  on the day of change were not statistically different between the infusion and control sessions ( $T_{8.7} = 0.44, p = 0.67$ ). Note that the surebet learning experiment was done before the free trial control experiment, so the lack of effect cannot be dismissed as a technical issue (i.e. infusions in the PPC were not effective at silencing activity). Overall, our results do not support the hypothesis that the PPC is required for shifting category boundaries: i.e. categorizing a lottery as being better or worse than the surebet.

### 3.4 Discussion

The neurobiology of decision-making under risk has been studied extensively in humans, non-human primates and rodents. However, there has been a gap in task design between the human-primate and rodent experiments, that most rodent studies focused on unexpected uncertainty where choices often reflected their sensitivity of reward history rather than risk attitudes. Here, we developed a risky choice task for rats, where animals made cue-guided decisions between a lottery and a surebet option on a trial-by-trial basis under expected uncertainty, as in human and primate experiments (but see Hocker et al., 2021, Constantinople et al., 2019b, for recent examples of rodent work on expected uncertainty). We developed the three-agent mixture model to decompose different elements of risk-preference, including  $\rho$  as exponent on the utility curve and  $\vec{\omega}$  as the weights for the rational, lottery and surebet agents. Modeling results showed that all subjects had decelerating, i.e. risk-averse, utility functions and their decisions were influenced mostly by the rational agent. We tested how inactivations of two cortical regions, the FOF and PPC, influenced choices in



**Figure 3.7.** Bilateral PPC inactivations did not impair learning of new surebet magnitudes. **A.** Schematic showing changing the surebet magnitude is equivalent to shifting the choice boundary. The data points were simulated from a risk-neutral agent using the three-agent model ( $\rho = 1, \sigma = 3, \omega_{rational} = 1$ ). A smaller surebet magnitude (light blue) horizontally shifts the psychometric curve leftwards, a larger surebet magnitude (dark blue) shifts the curve rightwards. The frequency-to-lottery mapping remains the same. **B.** Changing surebet magnitude from 6.8 to 3 shifted choices leftwards in one example animal. Combined trials from 6 sessions before the change are shown in gray, after the change shown in blue. One three-agent model was fit to all the trials and the parameters were used for ribbon extrapolation. **C.** Same as B but with  $0.6 \mu\text{g}$  per side bilateral PPC infusion, performed on the day of surebet change (from 3 to 6.8). **D.** The three-agent mixture model predicts the shifts in behavior well. One model was fit using all the sessions containing various surebet magnitudes for each animal. On x-axis is the predicted shift in probability choosing lottery ( $p(\text{Choose Lottery})$ ): the difference in  $p(\text{Choose Lottery})$  between model prediction using the new surebet magnitude and the session just before that change. On y-axis is the actual shift in  $p(\text{Choose Lottery})$ : the difference in  $p(\text{Choose Lottery})$  between the first session of a surebet change and the session before that change. Sessions with just surebet change are in blue ( $n = 21$ ; 4 animals), sessions with both surebet change and  $0.6 \mu\text{g}$  per side bilateral PPC infusions are in gold ( $n = 8$ ). The model prediction strongly correlated with the actual shift ( $R = 0.905, p < 0.001$ ). No significant difference in the actual shift was found between the blue and gold sessions ( $T_{8.7} = 0.44, p = 0.67$ ).

our task. These regions have been studied extensively in perceptual decision-making, but this is the first test of their causal role in economic decision-making. Bilateral FOF inactivations produced a profound bias towards the surebet, while bilateral PPC inactivations had minimal effect on the behavior. Model-based analyses of the results indicated that, without the FOF, subjects had utility functions that were substantially shifted towards risk-aversion. We constructed a biophysical model to show that inactivating the FOF node can produce outputs similar to the effects observed, providing a mechanistic explanation. Finally, we found that PPC was not causally involved in the learning of new categorization boundaries.

### 3.4.1 Role of FOF

Results from bilateral FOF inactivations show that FOF is an essential part of the circuitry underlying risky decision-making. Model-based analyses suggest that the change in behavior was likely due to a decrease in  $\rho$ , the curvature of the utility function,  $U = V^\rho$ . However, due to the small number of trials collected in bilateral FOF inactivation sessions, we cannot say definitively that the effect was exclusively on  $\rho$ ; it may have been a combination of both decreasing  $\rho$  and also increasing influence from the surebet agent on choice. Additionally, there are many functional forms of decision-making under uncertainty that we did not test (Chen and Stuphorn, 2018, Farashahi et al., 2019, Weber et al., 2004), which could lead to different interpretations. Nevertheless, the change in  $\rho$  is consistent with the finding that inactivation of monkey supplemental eye field reduced risky choices and the change was characterized by a decreased utility exponent (Chen and Stuphorn, 2018). Using a dynamical model, we demonstrated that a shift in  $\rho$  can be caused by a partial inactivation of

an action-value network whose activity guides choice (Samejima et al., 2005). This is similar to the theory that the FOF is part of a network for planning upcoming choice (Kopec et al., 2015). In fact, the interpretation of FOF activity encoding action-value is consistent with the previous interpretation (planning movement), since in perceptual tasks, only the correct side is rewarded, making it difficult to disentangle action-value from movement-planning. This action-value network may be the locus of transformation from value to utility, with the network properties (e.g. whether the gain of the network is greater or less than 1) determining whether animals are risk-seeking or risk-averse. One key difference between our current and previous findings, is that previously, sensory-guided choices (i.e. trials with no working-memory requirement) were not affected by silencing FOF (Erlich et al., 2015, Kopec et al., 2015). In Erlich et al. (2015), we posited that the FOF may be a bottleneck through which long-timescale integration of information could influence orienting decisions. Our results here suggest that the FOF may play a similar role for decisions that require integration of multiple attributes – in this case, lottery value and probability.

The idea that the FOF is part of an action-value network is largely consistent with the view that the FOF contributes to sensory-to-motor transformation (Scott et al., 2017, Insanally et al., 2019, Siniscalchi et al., 2016, Ebbesen et al., 2018, Barthas and Kwan, 2017), but reinterprets those findings as sensory-to-value transformations. Whether the same network also directly guides movement or transmits action-values to a downstream action-selection circuit (Essig et al., 2021, Shires et al., 2010) is an interesting question for future work: it has been previously demonstrated that changes in, e.g. excitatory drive, driven by a go-cue, can shift a network from monotonically encoding a task variable to encoding a binary choice (Machens, 2005). It is

important to note that, in our task, the surebet value is stable across trials and only the lottery needs to be evaluated on a trial-by-trial basis. We predict that in a task where the lottery is stable across trials and the surebet value varies trial-by-trial (and is indicated by a cue), silencing the FOF would shift animals away from selecting the surebet, whose value would require transformation on each trial. In tasks that require transformations for both surebet and lottery (e.g. Constantinople et al., 2019b), bilateral FOF silencing might result in increased decision noise.

Both left and right unilateral FOF inactivations led to a small bias towards leftward choices (FIGURE 3.2C), rather than an contralateral impairment, as previously reported (Erlich et al., 2015, Kopec et al., 2015, Hanks et al., 2015). However, in those studies, trials that did not require short-term memory were not biased by unilateral FOF inactivations, so the small effect is not particularly surprising. As 7 out of 8 animals had the surebet port on the left, the leftward bias can be interpreted as a weak bias towards choosing the surebet; i.e. a partial effect consistent with our bilateral silencing results.

### 3.4.2 Role of PPC

#### 3.4.2.1 Role of PPC in risky choice

Activity in PPC has long been associated with decision variables in economic choices. Platt and Glimcher (1999) first showed that activity in monkey lateral intraparietal cortex (LIP), a visuomotor area within PPC, is sensitive to expected reward magnitude and probability. Subsequently, Dorris and Glimcher (2004) found that neurons in monkey LIP encode relative subjective desirability of actions in a mixed-strategy game. Activity in human PPC also correlates with subjects' risk preferences

(Huettel et al., 2006). To date, we are not aware of any studies of the role of rodent PPC in economic decisions. Nonetheless, PPC encodes task-related variables during perceptual decisions (e.g. Hanks et al., 2015, Raposo et al., 2014, Goard et al., 2016). We were frankly disappointed that neither unilateral nor bilateral inactivation had any effect on the risky choices. As far as we are aware, this is the first experiment that directly tests the causal role of PPC in economic choices.

The null result is reminiscent of the null effects of PPC inactivation in the Poisson clicks task in rats (Erlich et al., 2015), and of LIP inactivation in the random dot task in monkeys (Katz et al., 2016). The null effect was unlikely the result of insufficient inactivation, as unilateral PPC infusions led to a significant ipsilateral bias in the free choice trials, where the decisions were guided by internal side preference rather than action value (FIGURE 3.3A). The free choice result replicates previous findings (Erlich et al., 2015, Katz et al., 2016) and is consistent with the literature on the role of rodent PPC in neglect (Bucci, 2009, Reep and Corwin, 2009), providing a clear positive control for the inactivation experiments. Taken together, our results demonstrate that PPC is not strictly necessary for making utility maximizing choices under risk.

It has been argued that rodent PPC is important for visually-guided decisions but not other modalities. For example, pharmacological inactivation of PPC has been shown to impair mice' performance in a visually-guided navigation task with a memory component (Harvey et al., 2012), and in a multi-sensory perceptual task but when only using visual but not auditory cues (Raposo et al., 2014). It was suspected that due to the anatomical proximity between PPC and the visual areas, these inactivation results may be caused by a muscimol spillover into the adjacent

visual cortex. However, recent experiments utilizing optogenetics have shown that, targeted inactivation of PPC during the stimulus period disrupted performance only on the visual but not auditory trials (Licata et al., 2017), and impaired decision sensitivity in a visually-guided task with variable delays in mice (Goard et al., 2016). As such, we cannot exclude the possibility that the null effect on risky choice may be due to the modality of stimuli used. However, silencing mouse PPC was shown to impair re-categorization of sounds in an auditory task (Zhong et al., 2019), so the controversy over the modality-specific role of PPC is not fully resolved.

### 3.4.2.2 Role of PPC in learning

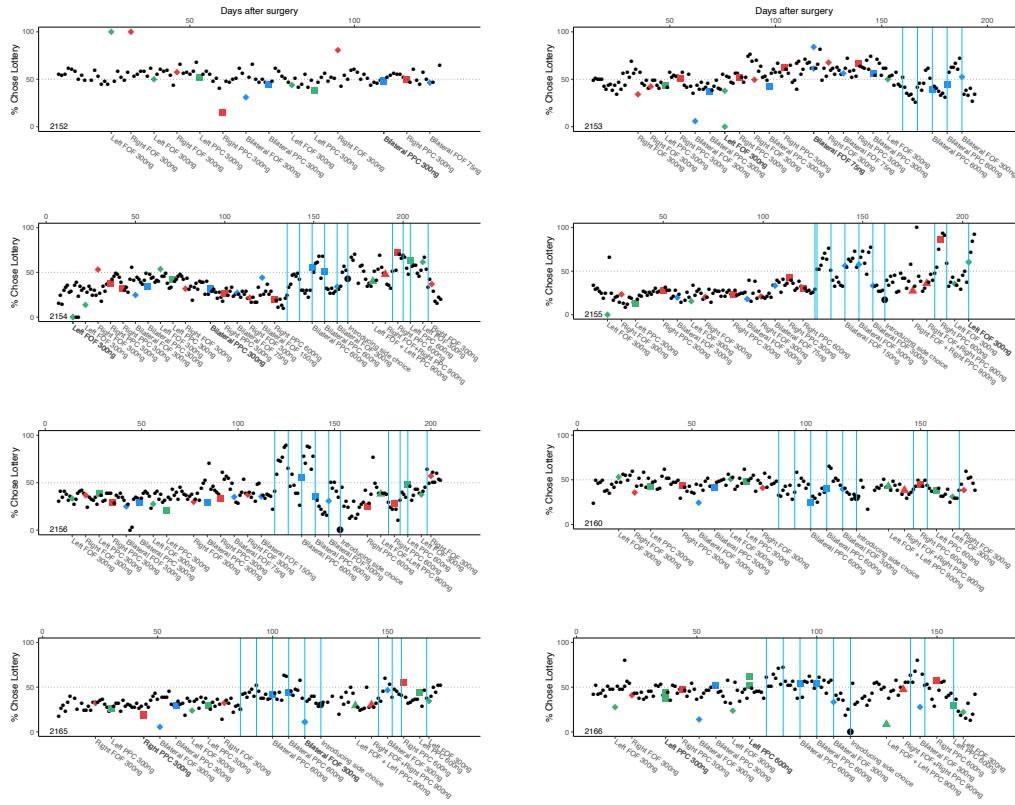
We have shown that inactivating PPC did not impair the animal's ability to shift their choices in response to changes in the value of the surebet, inconsistent with the findings from Zhong et al. (2019). There are some key differences in the design between our and their experiments that may explain this. First, their experiments were performed on head-fixed mice, whereas our rats were freely moving in the training box. Second, their mice had to categorize (or re-categorize) stimuli for the first time while PPC was inactivated. They never tested whether PPC was required for learning recurring shifts in the decision boundary. In contrast, the animals in our experiment were accustomed to changing surebet values for two weeks prior to PPC inactivation, understanding that the surebet value may change unexpectedly. Bucci and Chess (2005) found that PPC-lesioned rats had trouble learning the association between light and food if previously the light was presented without food. Interestingly, normal learning was observed in another cohort of PPC-lesioned rats that were not pre-exposed to the light. They attributed the impairment to PPC's role in directing

attention to the stimulus whose meaning surprisingly changes. If it is the case that PPC activity is required for the learning of ‘surprising’ shifts in existing associations, the discrepancy between our experiment and Zhong et al.’s can be then resolved.

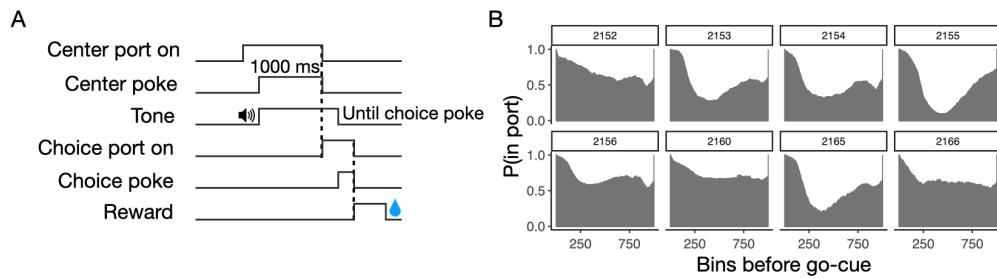
### 3.4.3 Conclusion

Studies of neurobiology of economic choice in rodents have mostly focused on the reward-valuation circuit: including the amygdala (Larkin et al., 2016, Orsini et al., 2017), basal ganglia (Stopper et al., 2013) and orbital-frontal cortex (Hocker et al., 2021, Constantinople et al., 2019a, Ogawa et al., 2013, Roesch et al., 2006, Gardner et al., 2017). Here, we examined the causal contribution of two cortical areas associated with planning orienting decisions, the FOF and PPC, whose analogous primate regions have been implicated in economic decision-making (Chen and Stuphorn, 2018, Platt and Glimcher, 1999). We found that FOF is a critical node in the circuit for decisions under risk, while PPC is not. Our results predict that FOF neurons participate in sensory-to-value transformation and would increase their activity monotonically with action-value.

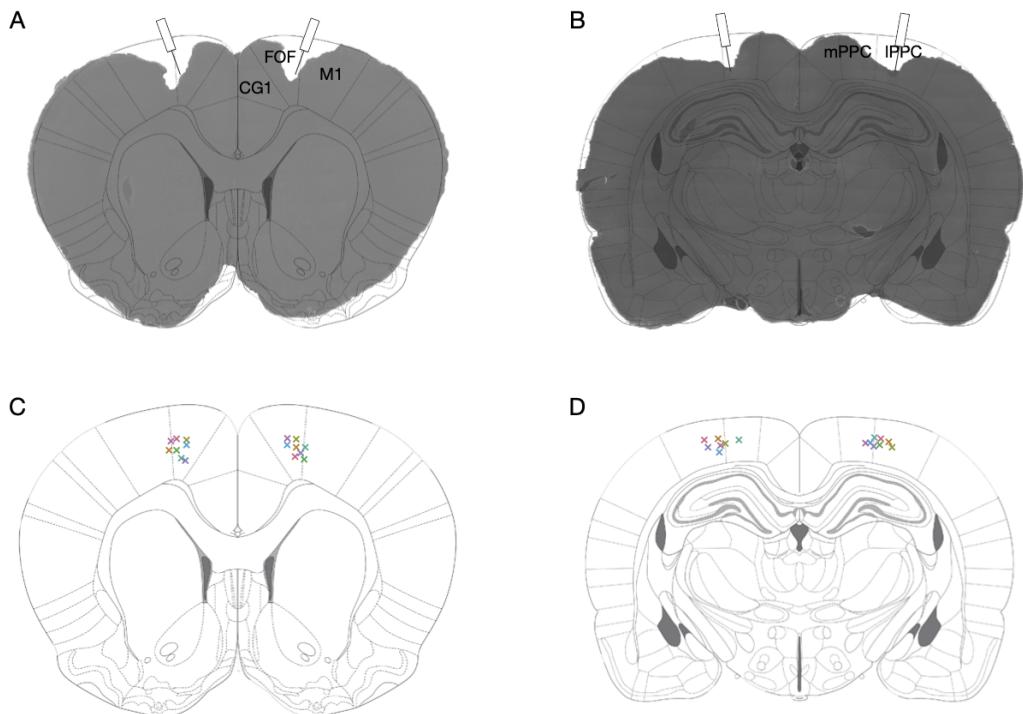
### 3.5 Supplementary figures



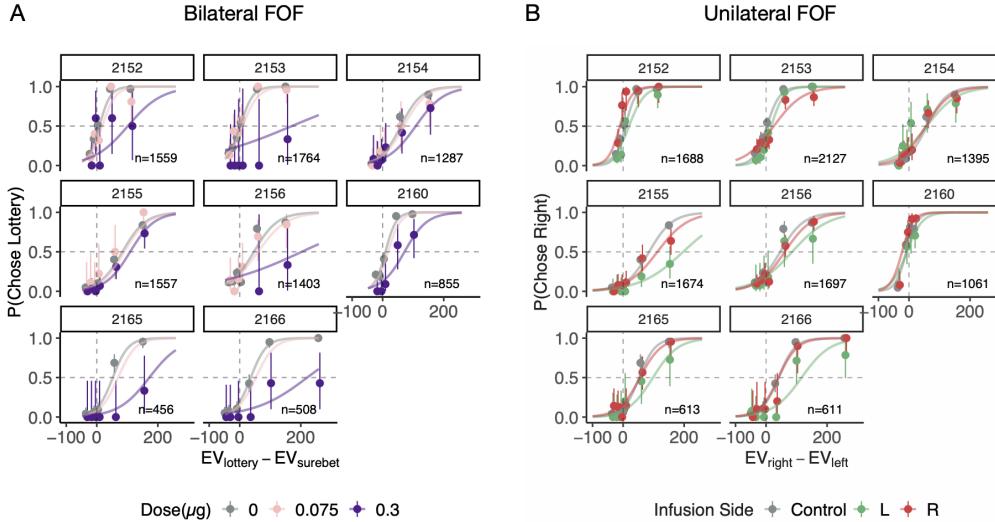
**Figure 3.8.** Timeline of percentage choosing the lottery in each session for each rat and a visual summary of the experimental treatments. Each point is the percentage choosing lottery for the given session. The number at the x-axis indicates the days passed since the surgical implantation of cannulae. Control days are shown as small black dots. Right infusions are shown in red, left infusions are in green, and bilateral infusions are shown in blue. FOF infusions are represented by diamonds, PPC infusions by squares, both FOF and PPC infusions by triangles. The blue bars indicate the day of a model-based surebet value change. The large black dot indicates the day when free choice trials were introduced. The bottom x-labels describe the details (side, region and dose) of each infusion.



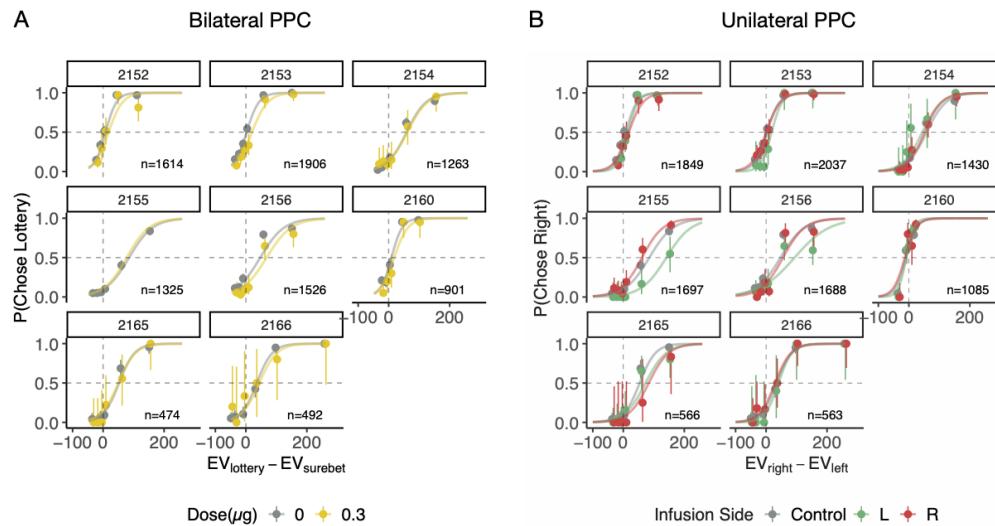
**Figure 3.9.** Task and behavior. **A.** Illustration of the timeline of a risky choice trial. **B.** The probability of being in the center port before the go-cue and after the first center port poke. For each trial, the period between the initial poke and the go-cue was segmented into 1000 bins. Go-cue is defined as the onset of the choice port lights. For each bin, a binary value was obtained to indicate whether the animal was in the port or not. The probability of being in the port for each bin was calculated by taking the mean of the binary vector. Only control trials were used for this analysis.



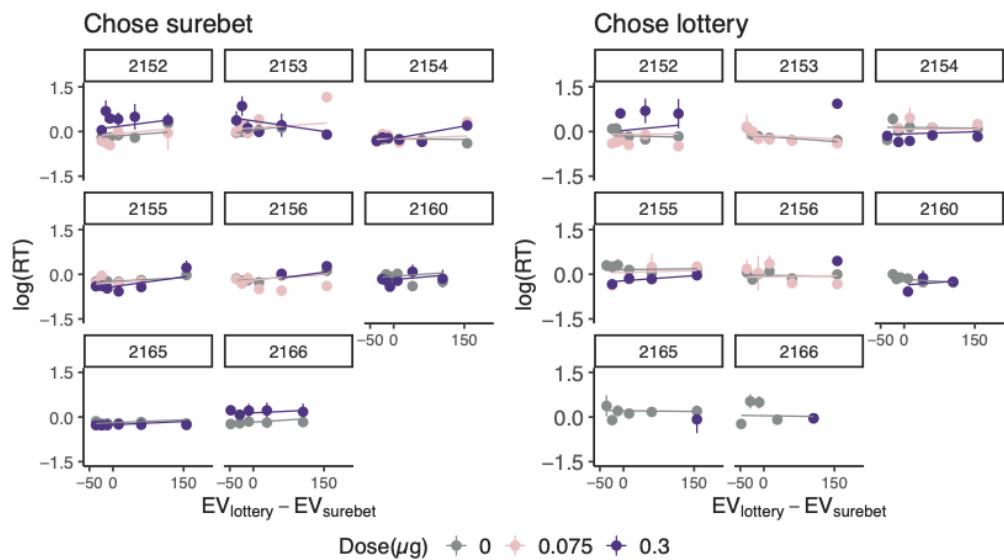
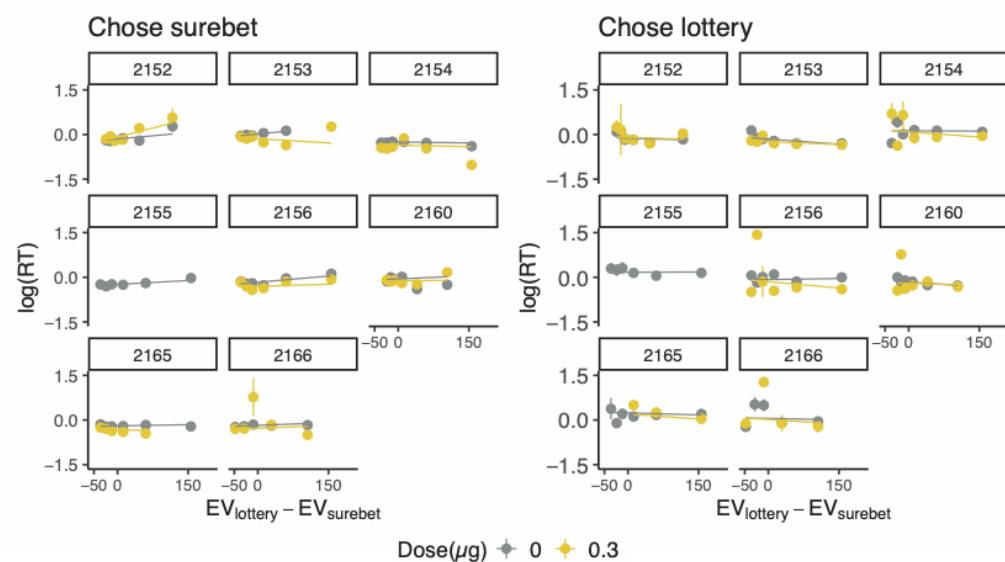
**Figure 3.10.** Histology. **A.** Coronal section of an example rat brain showing cannulae implanted at 20° in FOF, overlaid with a section 2.04 mm anterior to Bregma (Paxinos and Watson, 2004). Note, that in the nomenclature of Paxinos and Watson (2004) the area that we describe as the FOF is considered to be part of M2. CG1 = Cingulate Cortex. **B.** Coronal section of an example rat brain showing cannulae implanted at 10° in PPC, overlaid with a section 3.48 mm posterior to Bregma. mPPC = medial PPC, lPPC = lateral PPC. **C.** Actual cannulae placements in FOF, color represents the subject ID as in FIGURE 3.1E. **D.** Cannulae placements in PPC.



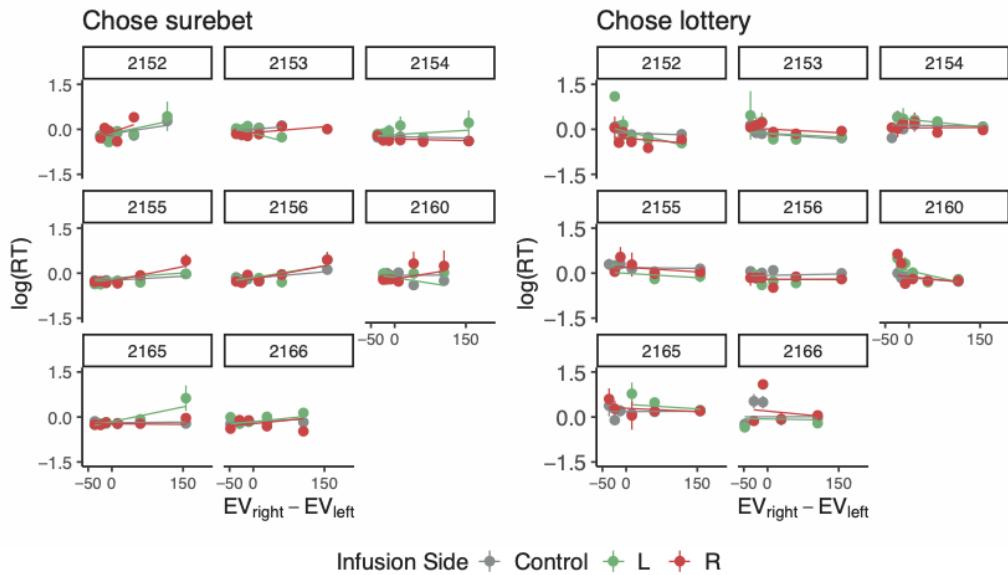
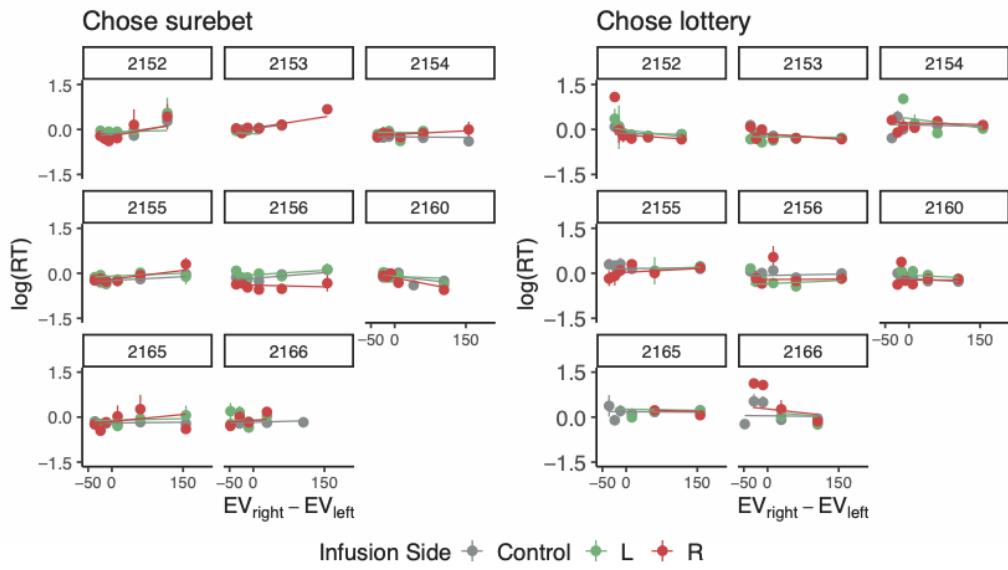
**Figure 3.11.** FOF inactivations. The circles with error bars are the binned mean and 95% binomial confidence intervals. The lines are the model predictions generated by the GLMM. **A.** Bilateral FOF inactivation with  $0.075\ \mu\text{g}$  and  $0.3\ \mu\text{g}$  muscimol per side. **B.** Unilateral FOF inactivation with  $0.3\ \mu\text{g}$  muscimol.



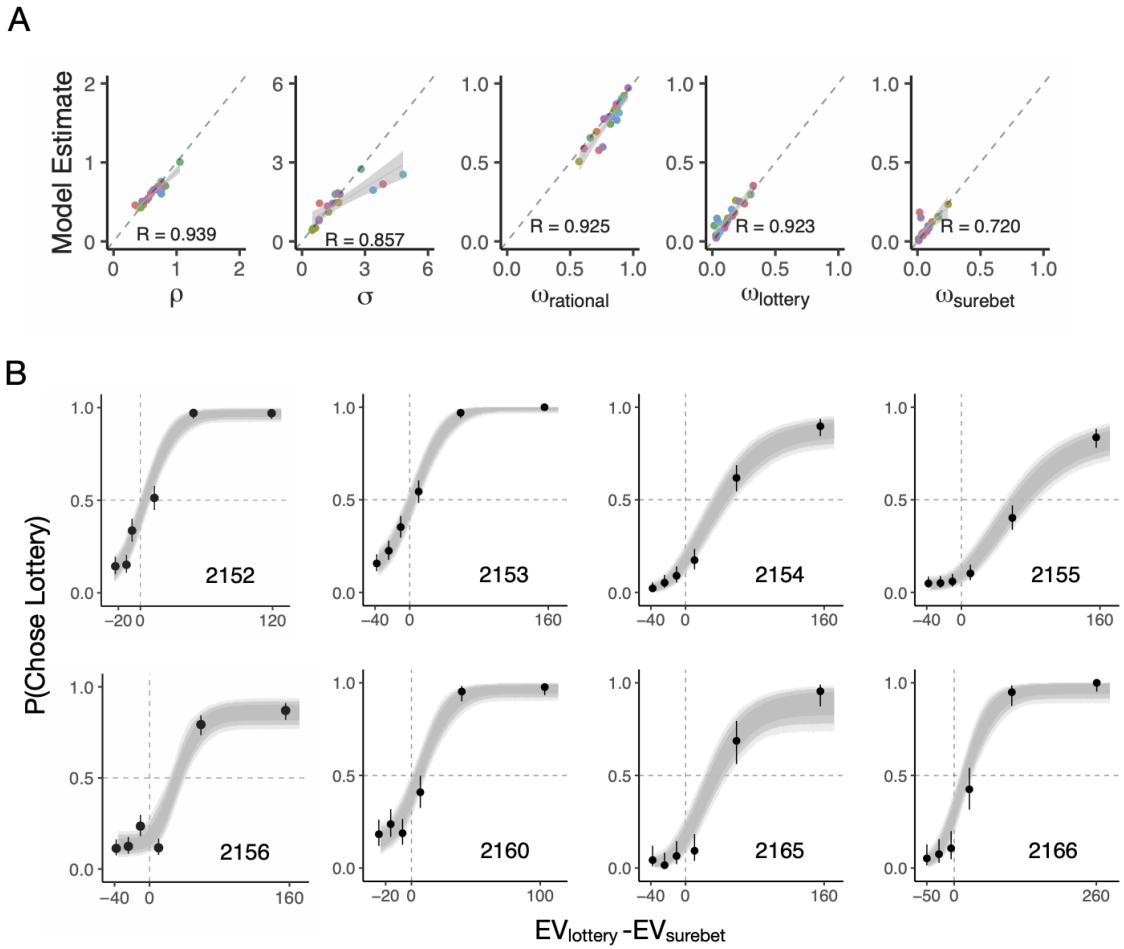
**Figure 3.12.** PPC inactivations. The circles with error bars are the binned mean and 95% binomial confidence intervals. The lines are the model predictions generated by the GLMM. **A.** Bilateral PPC inactivation with  $0.3\ \mu\text{g}$  muscimol per side. **B.** Unilateral PPC inactivation with  $0.3\ \mu\text{g}$  muscimol.

**A****Bilateral FOF****B****Bilateral PPC**

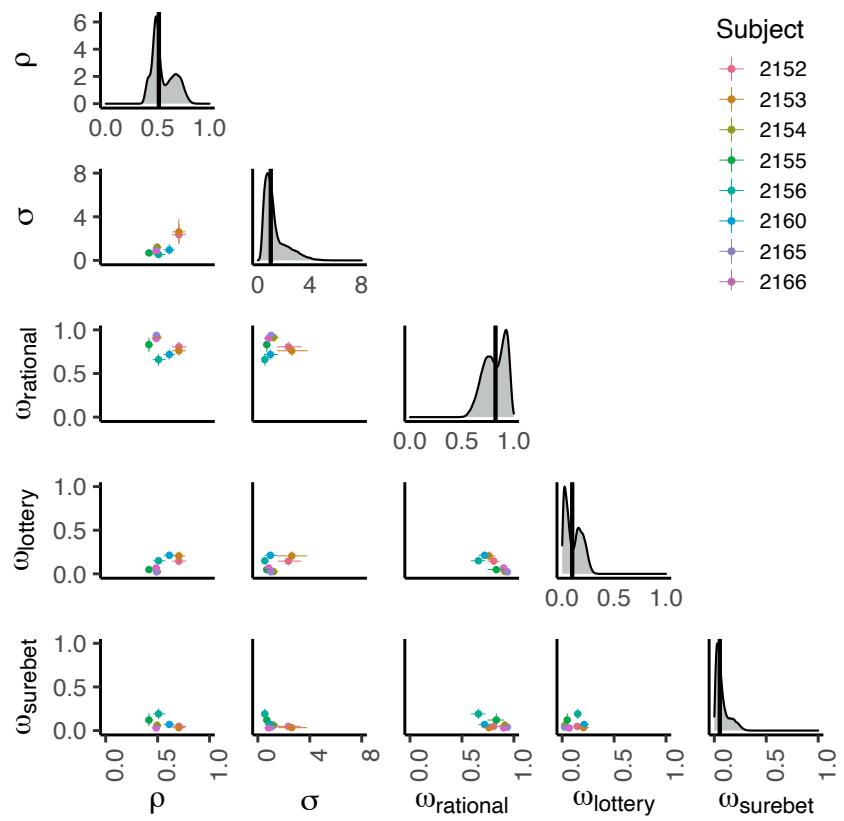
**Figure 3.13.** Reaction time (RT) and LMM model fits for bilateral FOF and PPC inactivations. The circles with error bars represent the mean and standard error of  $\log(\text{RT})$ . The lines are the model predictions generated by the LMM. **A.** Bilateral FOF inactivation. Reaction times are from the same trials as presented in FIGURE 3.11A. **B.** Bilateral PPC inactivation. Reaction times are from the same trials as presented in FIGURE 3.12A.

**A****Unilateral FOF****B****Unilateral PPC**

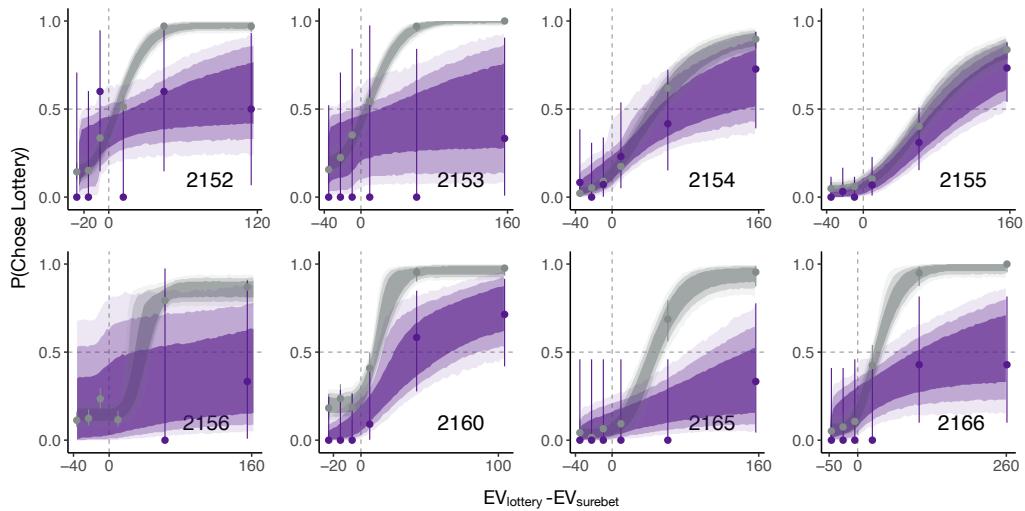
**Figure 3.14.** RT and LMM model fits for unilateral FOF and PPC inactivation trials. The circles with error bars represent the mean and standard error of log(RT). The lines are the model predictions generated by the LMM. **A.** Bilateral FOF inactivation. Reaction times are from the same trials as presented in FIGURE 3.11B. **B.** Bilateral PPC inactivation. Reaction times are from the same trials as presented in FIGURE 3.12B.



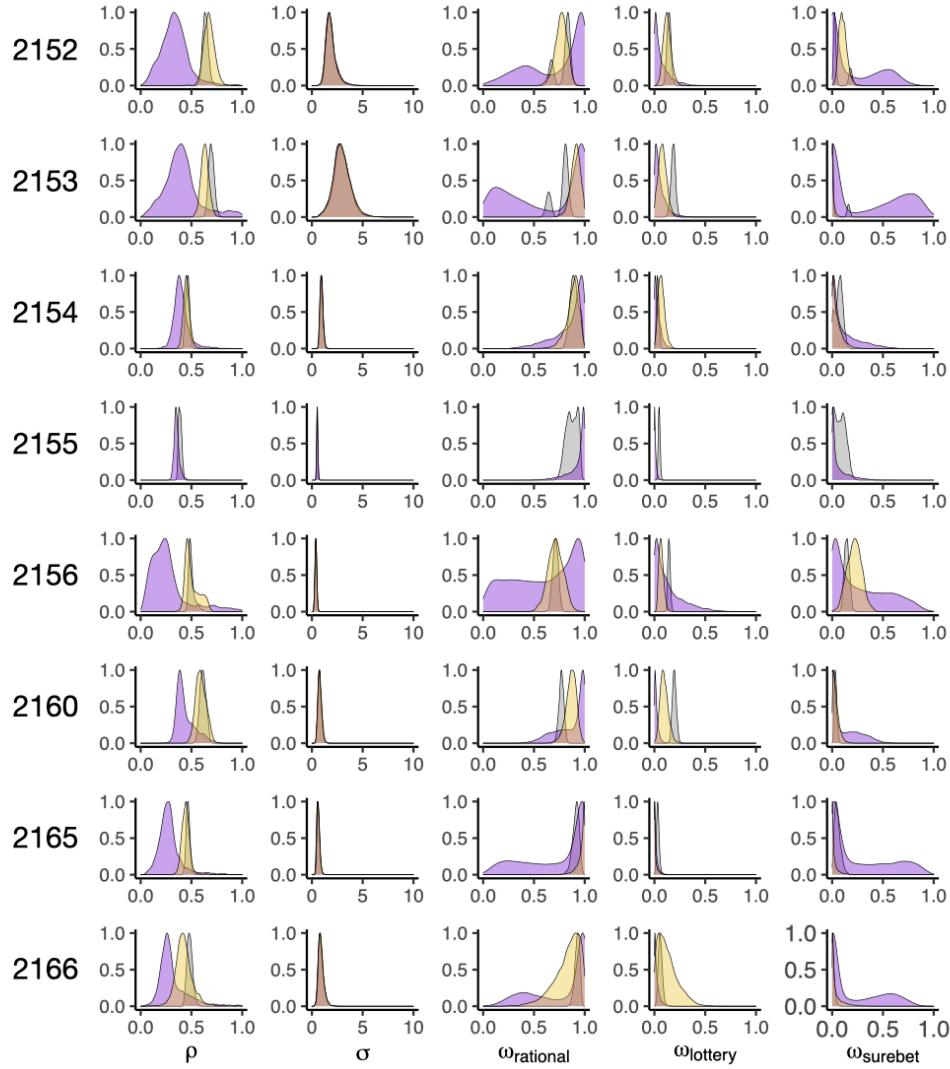
**Figure 3.15.** The three-agent mixture model. **A.** The model can recover the data-generating parameters well. Twenty Synthetic datasets were created by sampling from the same prior distributions as specified in Methods. The true parameter value is on the x-axis, the maximum a posteriori model estimation is on the y-axis. Color represents the identity of each synthetic dataset. All the parameters fall along the diagonal line. **B.** The psychometric data and model prediction from the three-agent mixture model for 8 animals. The circles with error bars are the binned mean and 95% binomial confidence intervals. The ribbons are model predictions generated using the fitted parameters. The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively. Data used are the same as the control sessions in FIGURE 3.1C.



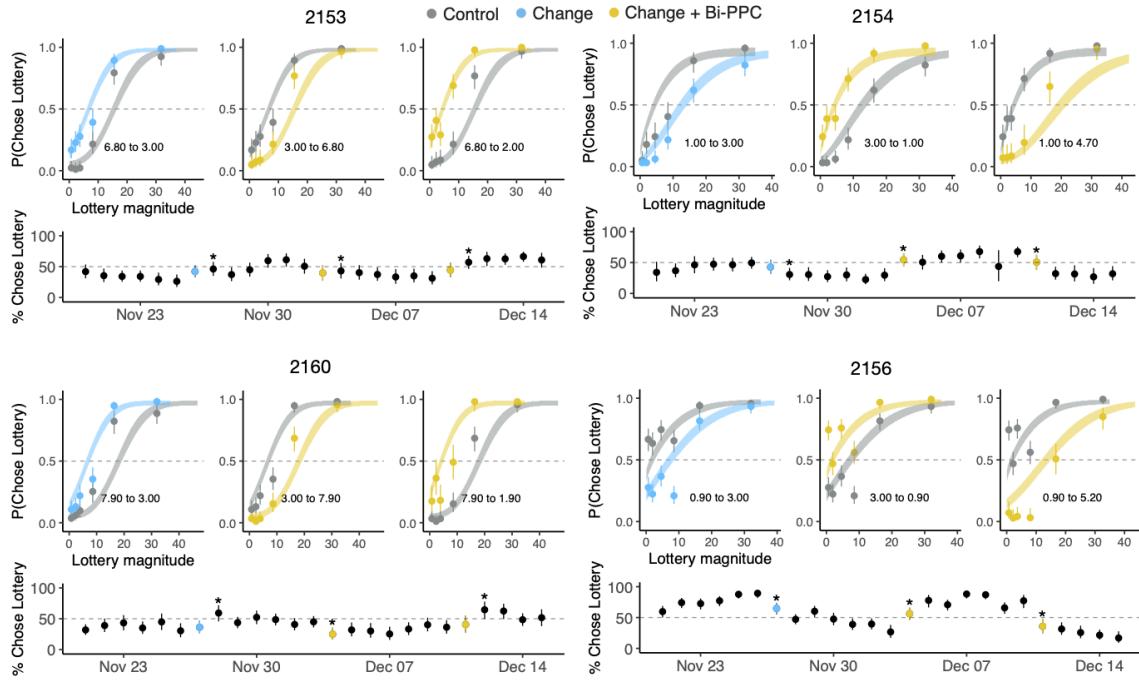
**Figure 3.16.** Summary of the fit model parameters from the control sessions of 8 animals. The mean and 95% confidence interval of each parameter pair are shown in the off-diagonal, colored by subject. Density plots of concatenated posterior samples for each parameter are on the diagonal, the black bar denotes the median.



**Figure 3.17.** Subjects' choices superimposed with the inactivation model fit on control (in gray) and bilateral FOF inactivation (in purple) dataset simultaneously. The circles with error bars are the binned mean and 95% binomial confidence intervals. The ribbons are model predictions generated using the fitted parameters. The dark, medium and light shade represent 80%, 95% and 99% confidence intervals, respectively.



**Figure 3.18.** Posterior distributions for each parameter using the inactivation model fit to the control (in gray),  $0.3 \mu\text{g}$  per side bilateral FOF inactivation (in purple) and  $0.3 \mu\text{g}$  per side bilateral PPC inactivation (in gold) dataset simultaneously. To allow easier visual comparison, all posteriors were normalized so that the peak of the distribution was set to 1. Since subject 2155 lost one PPC cannula, only the control and bilateral FOF fit was included here. From left to right:  $\rho$  is the exponent on the utility function,  $\sigma$  denotes the noise in utility representation,  $\omega_{rational}$  is the weight of the rational agent,  $\omega_{lottery}$  is the weight of the lottery agent, and  $\omega_{surebet}$  is the weight of the surebet agent.



**Figure 3.19.** Behavioral adaptation of subject 2153, 2154, 2156 and 2160 in the surebet learning experiment. Only one model was fit to all the trials and used for prediction for each animal. *Top three subpanels:* the circles with error bars are the binned mean and 95% binomial confidence intervals; the ribbons are generated using the fit parameter posterior of with 80% confidence intervals. Behavior from 6 sessions immediately before a surebet change is in gray, behavior from 7 sessions after a surebet change (including the very day) is in light blue if no infusion, in gold if with  $0.6 \mu\text{g}$  bilateral PPC infusion. Text annotation shows the old and new surebet magnitudes. *Bottom subpanel:* The percentage choosing lottery of each session. Asterisk indicates when change in choices can be significantly detected on that session compared to the previous 6 sessions with old surebet magnitude.

# **Chapter 4**

## **A rodent paradigm for studying perceptual decisions under asymmetric reward**

### **4.1 Introduction**

We are often required to make decisions based on noisy perceptual evidence where the costs associated with one choice are quite different than the costs associated with the other choice. Consider a dramatic example: a radiologist who needs to detect the presence (or absence) of a malignant tumor from a CT scan. The judgment itself should only be based on perceptual information, such as the density, shape and curvature of the anomalous cluster. However, the cost of mistakes in these medical decisions are highly asymmetric. If the tumor is present and the doctor says ‘no’ (miss), the patient will be discharged and possibly die from lack of treatment. On

the other hand, if the tumor is absent and the doctor says ‘yes’ (false alarm), the patient will just go through more tests until the determination of cancer is clear. When a perceptual judgment incurs asymmetric outcomes, the decision maker must integrate both the strength of sensory evidence *and* the distinct costs of mistakes, or different benefits of actions. In fact, such decisions are commonplace in the real-life, from judging if the yogurt has gone bad to an animal judging whether the noise is caused by a predator or a prey.

Laboratory tasks studying the cognitive and neural mechanisms of decisions under noisy sensory information typically assign equal reward to the options, where the subject is rewarded for categorizing the stimulus correctly. Recently, there is an emerging interest in studying perceptual decisions with asymmetric costs and reward, predominantly in humans (Diederich and Busemeyer, 2006, Diederich, 2008, Summerfield and Koechlin, 2010, Gao et al., 2011, Mulder et al., 2012) and non-human primates (Feng et al., 2009, Rorie et al., 2010). Despite the differences in tasks and models used, all of these experiments investigated how value information was integrated with sensory information during the decision process. One plausible mechanism is that the value information affects the processing of sensory information, such as by directly modulating the activity in primary sensory areas. Stănișor et al. (2013) found reward value is a good predictor of monkey V1 activity in a curve-tracing task, likely mediated by the top-down control of attention. Similar evidence was found in human V1, whose activity was modulated by reward value even in the absence of an overt saccade (Serences, 2008).

On the other hand, value information can also influence perceptual choices by adjusting the starting or ending point of the decision process. Evidence for this

view came from studies using variants of the drift diffusion model (DDM), which depicts the decision mechanism as a “diffusion” process, where the decision variable “drifts” towards a threshold based on upcoming sensory information (Ratcliff, 1978). Naturally, the model’s way to reflect an asymmetric starting point (as a result of value) would be to change the starting position for the decision variable. Such model-based analysis has shown that a shift in the starting position of DDM can best explain behavior in human subjects (Summerfield and Koechlin, 2010, Gao et al., 2011, Mulder et al., 2012) and non-human primates (Rorie et al., 2010).

These two alternative hypotheses, that value information exerts influence *on* or *separate from* sensory processing, predict that asymmetric reward should lead to differential neural activity in the sensory areas or secondary motor areas, respectively. Moreover, causal evidence for either hypothesis can be obtained by inactivating the candidate areas during the stimulus presentation or choice phase. The rat is an excellent model organism for studying the neurobiology of decision-making. Not only it is cost-effective, it also allows for manipulations with high temporal and spatial precision that are otherwise difficult in primates (e.g. Deisseroth, 2014, Kramer et al., 2013). Numerous groups have demonstrated that rats can learn complex perceptual and economic decision-making tasks, guided by visual and auditory cues (Constantinople et al., 2019b, Miller et al., 2017, Erlich et al., 2015, Zhu et al., 2021). Lak et al. (2020) trained mice on a task where they detected visual gratings with varying contrast, shown on the left or right monitor. The reward was asymmetric such that in alternating blocks, reporting one side correctly entailed a larger reward than the other side. To maximize reward, the animals must integrate reward history with trial-by-trial visual cues. This is the first rodent task, as far as we are aware of, that

investigates percept-value integration in a decision-making context. However, as the reward structure was not explicitly cued on each trial, this task is better suited for studying the learning of action-values than percept-value integration (Behrens et al., 2007). It is difficult to know exactly when in the trial the integration may be happening. Thus, we set out to develop a rodent task where the subject's choice is guided by both the perceptual and value-based components on a trial-by-trial basis.

## 4.2 Materials and methods

### 4.2.1 Subjects

Data from 7 male rats (4 Brown Norway, 3 Sprague Dawley; Vital River, Beijing, China) is included in this study. 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.

### 4.2.2 Behavior

#### 4.2.2.1 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) was fitted with 8 nose ports arranged in four rows (FIGURE 3.1A), 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 rats' interactions with the port. The port in the bottom row contained a stainless steel tube for delivering water reward. Animals were loaded and unloaded from the behavioral chambers by technicians daily on a fixed schedule. Each training session lasted for 90 minutes.

#### 4.2.2.2 The perceptual gambling 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 1 s – the ‘fixation’ period. As soon as the animal started fixation, a 500 ms tone would play from both speakers. The tone’s frequency (in  $\log_2(kHz)$  space) was sampled from a Gaussian distribution centered at 3 and truncated at 2 and 4, values corresponding to 8 kHz, 4 kHz and 16 kHz. Unless otherwise specified, we will use the  $\log_2(kHz)$  value throughout this manuscript. Specifically, the probability density function,  $\psi$ , describing the distribution of  $\log_2(kHz)$  tone frequencies was:

$$\psi(\mu, \sigma_s, \alpha, \beta; s) = \begin{cases} 0, & \text{if } s < \alpha \\ \mathcal{N} \sim (\mu, \sigma_s), & \text{if } \alpha \leq s \leq \beta \\ 0, & \text{if } s > \beta \end{cases} \quad (4.1)$$

where  $\sigma_s$  is the standard deviation and controls the difficulty of the perceptual task; it was tuned for each animal. The boundaries where  $\psi$  is truncated are defined by

$\alpha$  and  $\beta$ . The perceptual task required subjects to report whether the  $\log_2(kHz)$  of the tone was greater or less than 3. We counterbalanced the left / right assignment across animals, that animals with even subject ID were rewarded for tones  $\leq 3$  on the left, and animals with odd ID were rewarded for tones  $\leq 3$  on the right. We refer the correct port for frequencies lower than 3 as the ‘low port’, and the correct port for frequencies higher than 3 as the ‘high port’. After 1 s fixation, the animal was free to withdraw from the center port and poke into the left or right choice port. The animal was rewarded with the base amount if it chose correctly, no reward was delivered otherwise. If a trial had no flash and the animal was rewarded the base amount for choosing the correct port, we refer to these trials as ‘perceptual trials’. Around 30% to 65% of the total trials in a session were perceptual trials, the proportion was different for each animal (51.4 [30.7, 64.2], mean and 95% C.I.).

On some trials, concurrent with the tone, the three ports of one side would flash their yellow LEDs in the rate of 10 Hz, lasting for the entire duration of fixation. The selection of the flashing side was independent of the correct side indicated by tone frequency. If the flashing side coincided with the correct side, the animal would be rewarded with  $\kappa$  times of the *base reward* if it chose the correct port. The reward multiplier  $\kappa$  was tuned for each animal. If the flashing side was different than the correct side, the animal was rewarded the base amount if it chose the correct port. No reward was delivered for choosing the incorrect port. We refer to these trials as ‘perceptual gambling (PG) trials’. Around 35% to 70% of the total trials in a session were PG trials (48.5 [35.7, 69.2]). 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 300 s after trial start, or it did not make a choice 30 s after fixation. Violations were

excluded from all analyses.

#### 4.2.2.3 Training pipeline

Animal training took place in four distinct phases: the operant conditioning phase, the fixation phase, the perceptual phase and the perceptual gambling phase.

**The operant conditioning 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, randomly interleaved) before the reward port was illuminated for reward. They graduated to the next phase if they correctly performed these trials at least 40% of the session.

**The fixation phase** In the fixation phase, rats started by initiating the trial by poking into the center port. To facilitate initial learning, only two tones were presented (4 and 16 kHz) and the same tone was presented in blocks of 5 to 20 trials. The fixation duration started from 0 ms, and was increased by 5 ms every time the rat maintained fixation in the previous trial, otherwise it remained unchanged. Rats graduated to the next phase once the fixation time reached 1 s and they could reliably choose the correct port given the frequency (75% correct rate overall).

**The perceptual phase** The goal of the perceptual phase was to train the animals on the complete range of tone frequencies. Rats started with only 2 frequencies per

side ([2, 2.25] and [3.75, 4]) in blocks of 5 to 20 trials, more intermediate frequencies were added in pairs once they reliably differentiated the existing stimuli. The complete list of discrete frequencies was [2, 2.25, 2.5, 2.75, 3.25, 3.5, 3.75, 4], as they were spaced evenly apart in  $\log_2(kHz)$  space. Once the performance was stable on all discrete frequencies, we introduced continuous frequencies by sampling from one of two truncated Gaussian distributions:  $\psi(3, \sigma_s, 2, 3; s)$  and  $\psi(3, \sigma_s, 3, 4; s)$ , depending on the block. Initially,  $\sigma_s$  was set to be large to expose the animals to a wide range of frequencies and made the task relatively easy. Once the animals displayed sharp psychometric curves with continuous stimuli in blocks, we removed the block structure and sampled from the truncated Gaussian distribution  $\psi(3, \sigma_s, 2, 4; s)$ . The rats graduated from the perceptual phase if they showed reasonable psychometric curves on the continuous frequency range and understood that 3 was the decision boundary.

**The perceptual gambling phase** Rats entered the final perceptual gambling phase with good understanding of the frequency-to-choice mapping. The goal of this phase was to let animals learn the meaning of the light flash, which was introduced with a block structure. In a block of 20 to 30 trials, only one side would flash while the tone was still drawn from the truncated Gaussian distribution. In this phase, various task parameters were adapted to each animal's reward sensitivity to induce the 'perceptual gambling effect'. For example, if an animal did not shift its choice to the flashing side, we would increase  $\kappa$  to increase the expected value of the flashing side, decrease  $\sigma_s$  to make the trials more perceptually challenging and thus increase perceptual uncertainty, or increase the block length to help learning. The block length was gradually reduced to 1 - 3 once the subjects reliably shifted their

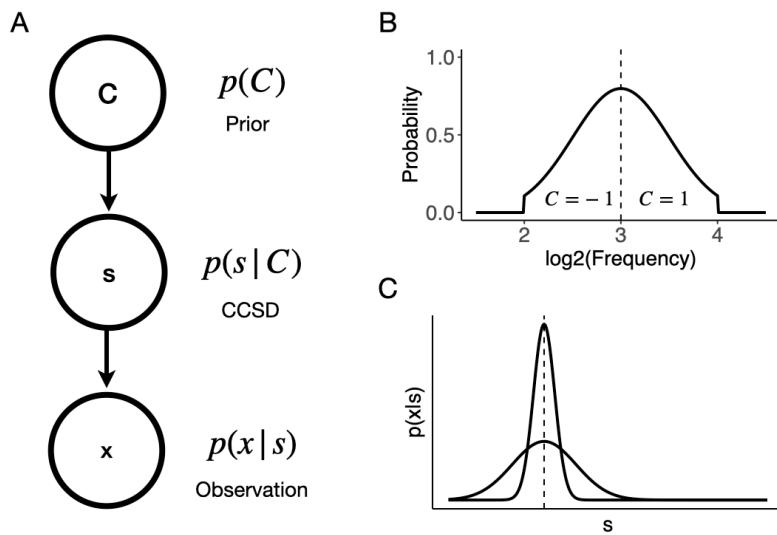
choices in response to the flashing side.

### 4.2.3 Modeling with Bayesian decision theory

The perceptual gambling task is a binary classification task with asymmetric action costs. On each trial, the rat has to take an action that requires inferring the correct class  $C$  of its auditory observation  $x$  from the actual stimulus frequency  $s$ . The probability of occurrence of each class is captured by the probability distribution  $p(C)$ , known as class priors. The distribution of the observation is specified conditioned on the class  $C$  and denoted by  $p(x|C)$ , this is known as the likelihood. In our task,  $p(x|C)$  cannot be directly known but had to be derived from other conditional distributions, which will be described below. Together, the distributions  $p(C)$  and  $p(x|C)$  define a ‘generative model’, a Bayesian description of how the observations arise from the auditory stimulus presented on each trial (Ma, 2019). The main assumption of Bayesian modeling is that the rat has learned the distributions specified in the generative model, and it utilizes this knowledge fully when inferring possible states of the world. This is done by using Bayes’ rule,

$$p(C|x) = \frac{p(x|C)p(C)}{p(x)} \quad (4.2)$$

where  $p(C|x)$  denotes the inferred posterior probability of a certain class given the stimulus frequency, and  $p(x)$  acts as a normalization factor. We describe our modeling process in three distinct steps: defining the generative model, computing the posterior distribution, and choosing an action to minimize cost (following Ma, 2019).



**Figure 4.1.** The generative model of Bayesian inference. **A.** Diagram of the generative model. The stimulus  $s$  is drawn from the class  $C$ , and the observation  $x$  is based on  $s$  only. **B.** Schematic of the mirror-imaged class-conditioned class distribution in our task. **C.** Schematic of the observation distribution  $p(x|s)$  given stimulus  $s$ . A higher perceptual noise (larger  $\sigma_p$ ) produced a wider distribution, whereas a lower perceptual noise (smaller  $\sigma_p$ ) gave a more concentrated distribution around  $s$ .

**Defining the generative model** The generative model is shown in FIGURE 4.1A. It has three nodes: the correct class  $C$ , the stimulus frequency  $s$ , and the animal's noisy observation  $x$ . The rat's goal was to correctly report  $C$ , making that the world state of interest.  $C$  can take on two possible values: 1 for the high port being correct, and -1 for the low port being correct. Associated with  $C$  is a distribution  $p(C)$ , which is specified by two values,  $p(C = 1)$  and  $p(C = -1)$ . In our task,  $p(C = 1) = p(C = -1) = 0.5$ , as the correct class was determined by the frequency relative to 3 drawn from the symmetrical Gaussian distribution. The stimulus distribution is thus a class-conditioned stimulus distribution (CCSD), which is denoted by  $p(s|C = -1)$  and  $p(s|C = 1)$  for the two classes, respectively. As the stimulus  $s$ , was drawn from the Gaussian  $\psi(3, \sigma_s, 2, 4; s)$  and then designated to be  $C = \text{sign}(s - 3)$ , the two CCSDs are mirror-images of each other (FIGURE 4.1B). Formally,

$$p(s|C = -1) = \begin{cases} \psi(3, \sigma_s, 2, 3; s), & \text{if } 2 \leq s < 3 \\ 0, & \text{otherwise} \end{cases} \quad (4.3)$$

$$p(s|C = 1) = \begin{cases} \psi(3, \sigma_s, 3, 4; s), & \text{if } 3 \leq s \leq 4 \\ 0, & \text{otherwise} \end{cases} \quad (4.4)$$

To complete the generative model, we need to define how the animals made observations based on the actual stimulus,  $s$ . Conventionally, the observation,  $x$ , is defined as a normal distribution centered at  $s$  with standard deviation  $\sigma_p$ , which denotes the perceptual noise of the animal.

$$p(x|s) \sim \mathcal{N}(s, \sigma_p) \quad (4.5)$$

**Computing the posterior distribution** Computing the posterior distribution  $p(C|x)$  involves both the class likelihood  $p(x|C)$  and class prior  $p(C)$ . The question is now how we can write the class likelihood in terms of the distributions specified in the generative model above. As observation  $x$  is only dependent on  $s$ , the class likelihood can be obtained by marginalizing over the intermediate variable  $s$ :

$$p(x|C) = \int p(x|s, C)p(s|C)ds \quad (4.6)$$

$$= \int p(x|s)p(s|C)ds \quad (4.7)$$

We can now write the posterior distribution as follows, where all the distributions have been specified:

$$p(C|x) = \frac{p(x|C)p(C)}{p(x)} \quad (4.8)$$

$$= \frac{\int p(x|s)p(s|C)ds}{p(x)} p(C). \quad (4.9)$$

**Choosing an action to minimize cost** If the task were just a perceptual task where the animals were rewarded equally for reporting the correct class, the optimal Bayesian decision maker should compare the two class posteriors and report the class

$\lambda(a C, v)$	$C = 1$			$C = -1$		
	$v = 1$	$v = -1$	$v = 0$	$v = 1$	$v = -1$	$v = 0$
$a = 1$	0	0	0	$r$	$\kappa r$	$r$
$a = -1$	$\kappa r$	$r$	$r$	0	0	0

**Table 4.1.** The action cost table.

with higher probability:

$$d = \log \frac{p(C = 1|x)}{p(C = -1|x)}, \quad (4.10)$$

where  $d$  is the decision variable and its sign indicates the chosen class. However, the key component of the perceptual gambling task was the asymmetric reward cued by flashing lights. According to BDT, the animal's decision should consider the uneven 'cost' of the actions and choose the one with minimal cost on each trial. We define the cost function as  $\lambda(a|C, v)$ , denoting the cost (loss of reward) of action  $a$  when the correct class is  $C$  and the flashing condition is  $v$ . Let  $v = 1$  denote flashes on the high side,  $v = -1$  for flashes on the low side, and  $v = 0$  denote no flashes at all. Further, let the base correct reward be  $r$  and the reward multiplier be  $\kappa$ . Then the cost function of the action  $a$  given the correct class  $C$  and flashing side  $v$  can be summarized in the following table:

For example, the action cost of reporting class 1 ( $a = 1$ ) when the true class is -1 ( $C = -1$ ) and the flashing side is also -1 ( $v = -1$ ) is  $\kappa r$ , as the animal is 'missing out' on  $\kappa r$  reward if it reported correctly. The action cost of reporting class 1 ( $a = 1$ ) when the true class is 1 ( $C = 1$ ) is 0 regardless of the flashing side, as it is the only action to be rewarded in this scenario. To incorporate the cost function into the decision

variable, we assume the animal has a representation of the posterior-weighted cost for each action. Concretely, the posterior-weighted cost for choosing class 1 becomes

$$d_1 = \lambda(a = 1|C = 1, v)^\rho p(C = 1|x) + \lambda(a = 1|C = -1, v)^\rho p(C = -1|x). \quad (4.11)$$

And for class -1 becomes

$$d_{-1} = \lambda(a = -1|C = 1, v)^\rho p(C = 1|x) + \lambda(a = -1|C = -1, v)^\rho p(C = -1|x), \quad (4.12)$$

where  $\rho$  is the exponent on the animal's utility function. In our task, we only tested 2 costs, so we could have, instead of an exponent, included a multiplicative reward scaling parameter. However, this is a classic functional form for marginal utility, that allows us to interpret animals with  $\rho < 1$  as risk averse and animals with  $\rho > 1$  as risk seeking. Finally, the decision variable can be expressed as the log ratio between  $d_1$  and  $d_{-1}$ :

$$d = \log \frac{d_1}{d_{-1}} \quad (4.13)$$

$$= \log \frac{\lambda(a = 1|C = 1, v)^\rho p(C = 1|x) + \lambda(a = 1|C = -1, v)^\rho p(C = -1|x)}{\lambda(a = -1|C = 1, v)^\rho p(C = 1|x) + \lambda(a = -1|C = -1, v)^\rho p(C = -1|x)}. \quad (4.14)$$

From the cost function table, we know that  $\lambda(a = 1|C = 1, v) = \lambda(a = -1|C = -1, v) = 0$ ,  $d$  thus becomes

$$d = \log \frac{\lambda(a=1|C=-1, v)^\rho p(C=-1|x)}{\lambda(a=-1|C=1, v)^\rho p(C=1|x)} \quad (4.15)$$

$$= \log \frac{\lambda(a=1|C=-1, v)^\rho}{\lambda(a=-1|C=1, v)^\rho} + \log \frac{p(C=-1|x)}{p(C=1|x)} \quad (4.16)$$

$$= \log \frac{\lambda(a=1|C=-1, v)^\rho}{\lambda(a=-1|C=1, v)^\rho} - \log \frac{p(C=1|x)}{p(C=-1|x)}. \quad (4.17)$$

Different from EQUATION 4.10, the sign of the decision variable  $d$  takes on the opposite value of the final class of choice. The reversion is due to the fact that the goal is to minimize the action cost rather than maximize the posterior distribution. Finally, we converted the decision variable into a probability of choosing  $C = 1$  using a *logistic* function:

$$p(\text{Choose } C = 1|\Theta, s, v) = \frac{1}{1 + e^d} \quad (4.18)$$

where  $\Theta$  refers to all the parameters in the model.

**The three-agent model** We observed that several animals exhibited ‘lapses’: poor performance even on very easy stimuli (Pisupati et al., 2019). In order to account for this behavior, we developed a three-agent model that includes a ‘rational’ agent that outputs  $p(\text{Choose } C = 1|\Theta, s, v)$  from the BDT model, and two stimulus-independent agents that either habitually choose the low or high port. The choice on each trial becomes a weighted outcome of the votes from three agents with their respective mixing weights  $\omega$ , each implementing a different behavioral strategy. Formally,

$$p(\text{Choose } C = 1 | \Theta, \vec{\omega}, s, v) = \vec{P} \cdot \vec{\omega} \quad (4.19)$$

$$= p(\text{Choose } C = 1 | \Theta, s, v) \cdot \omega_{rational} + 1 \cdot \omega_{high} + 0 \cdot \omega_{low} \quad (4.20)$$

$$\sum \vec{\omega} = 1 \quad (4.21)$$

The full model we used to fit animal behavior is thus a BDT-inspired hybrid model, we refer to it as the ‘mixture-BDT’ model. For notation simplicity, in the following sections we will use  $p_1$  to denote  $p(\text{Choose } C = 1 | \Theta, \vec{\omega}, s, v)$ .

#### 4.2.4 Analysis

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

##### 4.2.4.1 Generalized Linear (Mixed-Effects) Models

Generalized linear models (GLM) and generalized linear mixed-effects models (GLMM) were fit using the `stats` and `lme4` R packages (Bates et al., 2015). To test whether the animals were sensitive to both tone frequency and flashing side, we specified a mixed-effects model where the probability of choosing the high port was a *logistic* function of  $\log_2(kHz)$ , the flashing side and their interaction as fixed effects. The flashing side is a categorical variable with three levels: low side flash, high side

flash and no flash. The rat and an interaction of rat,  $\log_2(kHz)$  and the flashing side are modeled as within-subject random effects. In standard R formula syntax:

$$\text{choose\_high} \sim \log2\_kHz * \text{flash\_side} + (\log2\_kHz * \text{flash\_side}|\text{subjID}) \quad (4.22)$$

where `choose_high` is 1 if the high port was chosen and 0 if the low port was chosen; `subjID` is the subject ID for each rat.

To test whether an individual animal was sensitive to tone frequency and flashing side with each  $\sigma_s$  and  $\kappa$  combination, we specified a GLM as follows:

$$\text{choose\_high} \sim \log2\_kHz * \text{flash\_side} \quad (4.23)$$

Only the  $\sigma_s$  and  $\kappa$  combination that resulted in a significant main effect of `flash_side` was included in this study (see TABLE 4.2).

To test whether the outcome of the previous trial affected choice on the current trial, we first classified the previous trial's outcome into four categories: the animal chose the low port and was rewarded, the animal chose the low port and was unrewarded, the animal chose the high port and was rewarded, and the animal chose the high port and was unrewarded. A GLMM was specified:

$$\text{choose\_high} \sim \log2\_kHz + \text{prev\_outcome} + (\log2\_kHz + \text{prev\_outcome}|\text{subjID}) \quad (4.24)$$

where `prev_outcome` is a categorical variable with four levels as described above.

To test whether the animal has a tendency to repeat its previous choice, we specified a GLMM as follows:

$$\text{choose\_high} \sim \log2\_kHz + \text{prev\_choice} + (\log2\_kHz + \text{prev\_choice} | \text{subjID}) \quad (4.25)$$

where `prev_choice` is 1 if the high port was chosen on the previous trial and 0 if the low port was chosen.

#### 4.2.4.2 Trial difficulty analysis

To understand how perceptual difficulty affected the animal's shift towards the flashing side, we employed a model-based analysis. First, we obtained the animal's perceptual sensitivity  $\sigma_p$  using the aforementioned mixture-BDT model. Then, we computed Z-score for each tone frequency  $s$  presented to this animal using the formula:  $Z = (s - 3)/\sigma_p$ . Based on the Z-score, the middle 33% trials ( $-0.426 \leq Z < 0.426$ ) are labeled as 'Hard' trials, the 16.5% left and right to the hard trials are labeled as 'Medium' trials ( $-0.95 \leq Z < -0.426$ ;  $0.426 \leq Z < 0.95$ ), and the 16.5% left and right to the medium trials are labeled as 'Easy' trials ( $Z < -0.95$ ;  $Z \geq 0.95$ ). By dividing trials this way, we ensured equal proportions of easy, medium and hard trials while taking into account the animal's perceptual sensitivity. After computing the absolute change in percentage choosing the high port induced by light flashes for each difficulty condition, we performed a linear mixed-effects model (LMM) to test significance:

$$\text{delta} \sim \text{difficulty} + (\text{difficulty}|\text{subjid}) \quad (4.26)$$

where `delta` refers to the absolute change.

#### 4.2.4.3 Model 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. The prior over the utility exponent  $\rho$  was  $\text{Lognormal}(\log(1), 0.3)$ , a weakly informative prior that prefers  $\rho$  to be to risk-neutral. The prior over perceptual noise  $\sigma_p$  was  $\text{Lognormal}(\log(0.3), 0.1)$ , a reasonable range in  $\log_2(kHz)$  space. The prior over the mixing weights  $\vec{\omega}$  was a Dirichlet distribution with the concentration parameter  $\alpha = [6, 2, 2]$ . The resulting  $\omega_{rational}$  distribution was broad and had the mean of 0.6, both  $\omega_{high}$  and  $\omega_{low}$  distribution had the mean of 0.2. By attributing more weight to the rational agent over the habitual agents, the prior reflected our selection of the experimental animals - only animals whose choices depended on the auditory cue were included. Four Markov chains with 1000 samples each were obtained for each model parameter after 1000 warm-up samples. The  $\hat{R}$  convergence diagnostic for each parameter was close to 1, indicating the chains mixed well.

#### 4.2.4.4 Sigmoid function

The four-parameter sigmoid function was specified as follows:

$$y = w_2(1 - w_1) + \frac{w_1}{1 + e^{-b(x-x_0)}} \quad (4.27)$$

where  $x$  is the tone frequency in ,  $y$  is the probability choosing high port, and the four parameters are:  $x_0$ , the inflection point of the sigmoid, controlling horizontal shifts;  $b$ , the slope of the sigmoid;  $1 - w_1$ , the total lapse rate, and  $w_2$ , representing the fraction of lapses that are low to high lapses. The sigmoid model was fit individually to each flash condition in each subject's dataset using Stan.

#### 4.2.4.5 Synthetic datasets

To test the validity of the mixture-BDT model, we first created synthetic datasets with parameters generated from the prior distributions described above. The model was used to fit on the synthetic datasets, and was able to recover the generative parameters accurately (FIGURE 3.15). This assured that the model had no systematic bias in estimating the parameters.

#### 4.2.4.6 Model prediction confidence intervals

To estimate the confidence intervals with model prediction as in FIGURE 3.4A, we first generated a synthetic dataset with regularly spaced sound frequencies (incremented by 0.01). After parameter sampling in each iteration (in the `generated quantities` block), the sampled parameters were used to predict the choices given the

synthetic offers. The resulting output is a `n_iter` × `n_sound` matrix, where `n_iter` is the number of iterations and `n_sound` is the length of unique stimulus frequencies. Finally, ±80, ±95 and ±99 confidence intervals for each offer were estimated by taking the respective quantiles of the `n_iter` predicted choices.

#### 4.2.4.7 Mixture-BDT optimality analysis

To understand the relationship between  $\rho$  and  $\sigma_p$  in obtaining maximum possible reward, we first created a synthetic task dataset with 1000 trials. The tone frequency was drawn from a truncated Gaussian centered at 3 and a standard deviation of 0.6. The  $\kappa$  was set to 5. There were equal proportions of the high flash, low flash and no flash trials ( $\sim 33\%$ ). We then created a mixture-BDT agent that is fully rational ( $\vec{\omega} = [1, 0, 0]$ ). A grid search was performed to find the total reward for each combination of  $\rho$  (0 to 1.5, incremented by 0.1) and  $\sigma_p$  (0 to 0.5, incremented by 0.1). The total reward ( $R$ ) was computed as follows:

$$R = \sum_i^{1000} (1 - p_1^i) \cdot r_{low}^i + p_1^i \cdot r_{high}^i \quad (4.28)$$

where  $p_1^i$  is the probability of choosing the high port from the mixture-BDT agent on the  $i$ -th trial,  $r_{low}^i$  is the reward delivered if choosing the low port on the  $i$ -th trial, and  $r_{high}^i$  is the reward delivered if choosing the high port.

## 4.3 Results

### 4.3.1 The perceptual gambling task

To establish a rodent framework to study decisions guided by both perceptual and value cues, we developed the perceptual gambling task. It was named so because although the correct decision was only informed by the perceptual cue, a reward-maximizing subject would choose the side with larger reward when the perceptual evidence was weak, effectively ‘gambling’ for more reward (FIGURE 4.2A). For example, imagine the subject was 75% certain that the stimulus should be categorized as ‘high’. But, on this trial, the subject knew that a correct ‘high’ response would be rewarded with 1 drop of water, while a correct ‘low’ response would fetch 8 drops. Then, the expected value of responding high would be  $P_{high} \cdot V_{high} = 0.75 \cdot 1 = 0.75$ . The expected value of responding low would be  $P_{low} \cdot V_{low} = (1 - 0.75) \cdot 8 = 2$ . So, this task asks the animal to gamble based on its perceptual confidence (which is experimentally varied by requiring subjects to make easy and difficult perceptual decisions) and the values of the two responses.

Subjects were first trained on the pure perceptual version of the task, with symmetric rewards, and we refer to these trials as ‘perceptual’ trials. On each trial, after self-initiation by poking into the center port, subjects fixated for 1 s while a tone would play from both speakers, and its frequency relative to 8 kHz indicated whether the left or right port was correct (counter-balanced across animals). We refer the correct port for frequencies lower than 8 kHz as ‘low port’, and the correct port for frequencies higher than 8 kHz as ‘high port’. The tone frequency (in  $\log_2(kHz)$  space) was drawn from a truncated Gaussian distribution  $\psi(3, \sigma_s, 2, 4; s)$ , where  $\sigma_s$  is

the standard deviation and controls the difficulty of the perceptual task. The smaller  $\sigma_s$  is, the more concentrated the auditory stimulus is around the decision boundary, and the more perceptually challenging the trials will be (FIGURE 4.2C). Once the animals showed good performance on the perceptual trials, we introduced the value cues during fixation by flashing the yellow LEDs of the three left or right ports (FIGURE 4.2A). The choice of the flashing side was independent of the tone frequency. We delivered the perceptual and value cues through the auditory and visual modality, respectively, to avoid any effects from intra-modality attention. For example, a louder tone from one side might interact with the animal's judgment of its frequency. If the correct port was on the same side as the flashing ports, correct responses resulted in a large reward (base reward  $\times \kappa$ ). Alternatively, if the flashing side was the incorrect side, the animal was only rewarded the base amount for choosing the correct port. These trials are referred as 'perceptual gambling (PG)' trials. The perceptual trials were randomly interleaved with PG trials in a session, the ratio of the two trial types was different for each animal.

Training of the task was difficult, as the animal's performance was highly sensitive to task parameter values, especially the difficulty of the perceptual task ( $\sigma_s$ ) and reward asymmetry ( $\kappa$ ). This is not surprising, Kepecs et al. (2008) trained rats on an odor discrimination task and the animals were only rewarded (nor not) after a variable delay. While it was waiting for the reward, the rat had an option to 're-initiate' the trial by leaving the choice port and start again, the frequency of which should correlate with its confidence of the perceptual decision. It was later reported that the training was also very parameter-sensitive, as the reward delay interacted with the rat's temporal discounting function, similar to how the perceptual difficulty interacted

with the reward sensitivity in our task (Kepecs and Mainen, 2012). Nonetheless, we successfully trained 7 animals on the task (see TABLE 4.2 for the task parameters specific to each animal). These 7 subjects were sensitive to both the auditory cue and the flashes with at least one set of task parameters, quantified by a generalized linear model (GLM). When the flash was first introduced, the animals did not show any bias towards the flashing side. Thus, the shifts in choices caused by the flashes were learning-induced, rather than bottom-up-attention-induced changes of behavior.

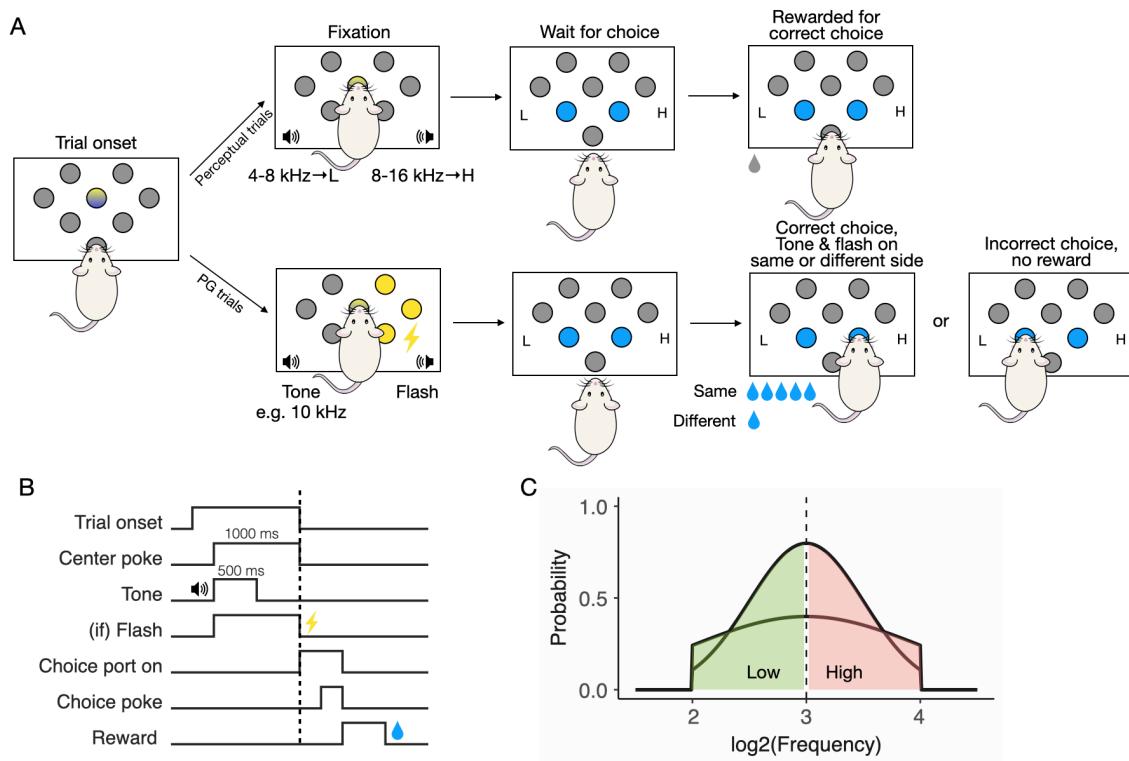
	2077	2078	2083	2085	2109	2124	2143
$\sigma_s$	1	1	1	1	0.3	0.3	0.3
$\kappa$	5	5	5	5	25	20	15

**Table 4.2.** The perceptual difficulty ( $\sigma_s$ ) in  $\log_2(kHz)$  space and reward multiplier of the flashing side ( $\kappa$ ) effective for each animal.

### 4.3.2 Behavior

We trained 7 male rats (4 Brown Norway, 3 Sprague Dawley) on the perceptual gambling task. As expected, the animals performance was a function of evidence strength (see an example animal in FIGURE 4.3A; see population aggregates in FIGURE 4.3C). Moreover, the animals reliably shifted their choices towards the side with flashing lights (see example shift in FIGURE 4.3B, all animals in FIGURE 4.8). These effects were quantified using a generalized-linear mixed-effects model (GLMM). There was a significant main effect of tone frequency ( $\beta_{freq} = 2.86 \pm 0.04, p < 0.001$ ), and a significant interaction between tone frequency and flash on the low side ( $\beta_{freq:side} = -0.13 \pm 0.15, p < 0.05$ ). Interestingly, flashing on the high side did not affect behavior significantly on the group level (all  $p > 0.05$ ).

The premise of the task is that the animal should prefer the flashing side more

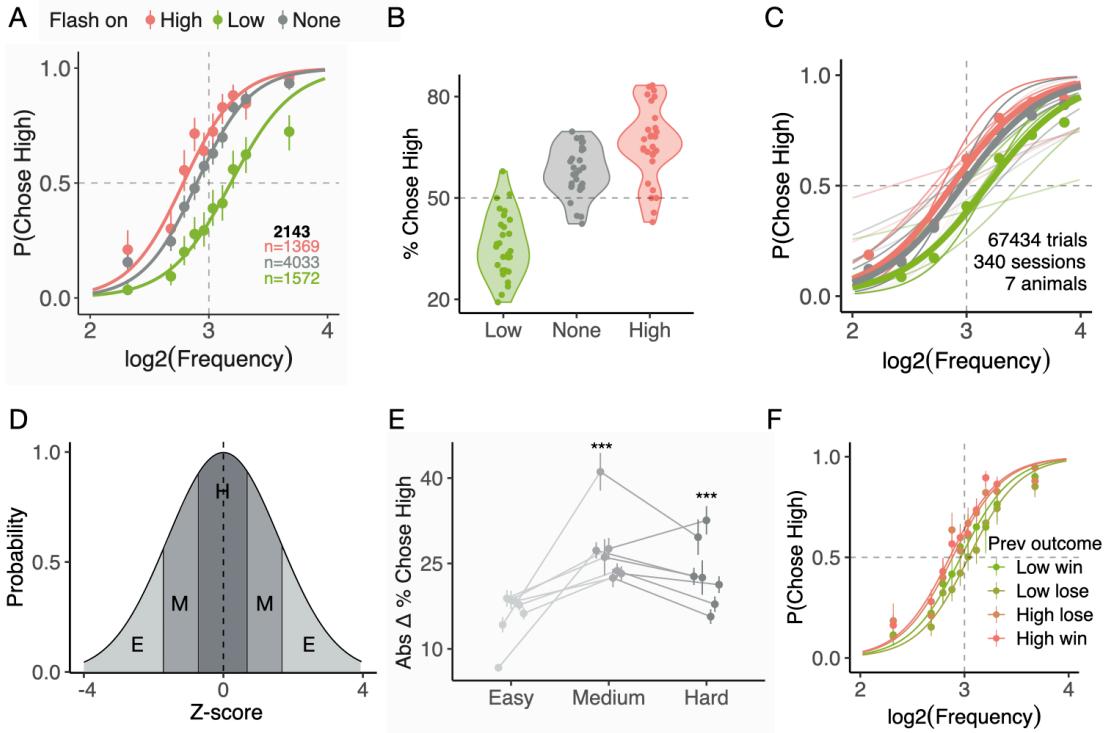


**Figure 4.2.** The perceptual gambling task. **A.** Schematic of the perceptual gambling task. Each trial began with the onset of a central LED, which cued the animal to poke into the center port and hold there for 1 s. A tone was played for 500 ms, its frequency relative to 8 kHz indicated whether the left or right port was correct. After 1 s, the animal withdrew from the center port and made a choice poke into the left or right port, and was rewarded with a small amount of water if chose correctly. Around 60% of the trials were pure perceptual trials, although the proportion varied across animals. The remaining trials had one side flashing yellow LEDs simultaneously with the tone, the light flashes would last for 1 s. If the flashing side was the same side as the correct side, choosing correctly would result in a large reward. If the flashing side was the different side than the correct side, choosing correctly would result in the same base reward. See more detailed task description in Methods. **B.** Schematic of the trial structure. **C.** The tone frequency was drawn from a truncated Gaussian,  $\psi(3, \sigma_s, 2, 4; s)$ , where  $\sigma_s$  was tuned for each animal in different training stages, but was otherwise fixed within sessions. A larger  $\sigma_s$  results in a wider distribution and easier trials, and a smaller  $\sigma_s$  results in a distribution more concentrated around the decision boundary.

when the stimulus was close to the decision boundary. In other words, when perceptual evidence was weak, the value information should have more influence on choice. To test whether this was true, we divided trials into easy, medium and hard trials based on their evidence strength in relation to the animal's perceptual noise, which was estimated using a Bayesian model (see description in the next section; FIGURE 4.3D). Using a linear mixed-effects model (LMM), we found that trial difficulty significantly affected the absolute shift in percentage choosing the high port ( $\beta_{hard} = 0.07 \pm 0.03, p < 0.001$ ;  $\beta_{medium} = 0.11 \pm 0.03, p < 0.001$ , FIGURE 4.3E). This is in line with our prediction that the subjects should shift their choices more for medium and hard than easy trials, although the reason why they shifted more for medium than hard trials was unclear. Finally, we found that the choices on the current trial were significantly influenced by the outcome of the previous trial (GLMM, all  $p < 0.001$ ; see an example in FIGURE 4.3F, see all animals in FIGURE 4.9). Although the individual history effects differ, overall, the animals had a tendency to repeat its previous choice ( $\beta_{prev\_choice} = 0.60 \pm 0.05, p < 0.001$ ).

### 4.3.3 A three-agent mixture model with Bayesian decision theory

While the GLMM results indicated that the animal's choices were sensitive to both perceptual and value cues, it does not provide insight into the cognitive processes underlying task performance. To better understand how our animals integrated perceptual and value information, we developed a Bayesian decision theory (BDT) model (following Ma, 2019). Bayesian modeling starts with a generative model, specifying how the subject's observation come about given the statistics of the environ-



**Figure 4.3.** Animal behavior in the perceptual gambling task. **A.** Example subject performance from 28 sessions with  $\sigma_s = 0.3$  and  $\kappa = 15$ . The probability of choosing the port associated with high frequencies is plotted as a function of  $\log_2(kHz)$ , where 3 was the decision boundary. The circles with error bars are the mean  $\pm 95\%$  binomial confidence intervals. The lines are the psychometric curves generated by a generalized linear model. The colors represent the flashing condition, with red - high side flash, green - low side flash, gray - no flash. **B.** Violin plot showing distribution of the percentage choosing the high port under three flashing conditions. Data used here is the same as in A. Each dot represents the mean percentage choosing the high port from this condition in each session. **C.** Performance from all 7 subjects trained on the task. The thick colored lines are the psychometric curves generated by a generalized linear model fit to all the sessions together. The thin colored lines are from a generalized linear model fit to each subject's dataset individually. **D.** Trials were divided into hard (H, hardest 33% of trials), medium (M, next hardest 33% of trials) and easy (E, easiest 33% of trials) based on each animal's perceptual noise. See details in Methods. **E.** The absolute change in percentage choosing the high port induced by light flashes in easy, medium and hard trials. The changes induced by either low and high flashes were averaged. Animals significantly shifted more in medium than easy trials ( $\beta_{medium} = 0.11 \pm 0.03, p < 0.001$ ), and more in hard than easy trials ( $\beta_{hard} = 0.07 \pm 0.03, p < 0.001$ ). **F.** 2143's choice on the current trial was influenced by the outcome of the previous trial.

ment, which is usually set by the experimenters. Using the Bayes' rule, the subject then combines its prior with the observation to obtain the posterior, a probability distribution reflecting both the observed measurement and its prior belief. Finally, a Bayesian decision maker chooses an action in a principled manner by minimizing a cost function  $C(s, a)$ , which is determined by the state of the world  $s$  and the action  $a$ . The BDT framework is well suited for our task, as the animal acts by integrating a noisy perceptual stimulus (observation) and asymmetric reward associated with each choice (cost function).

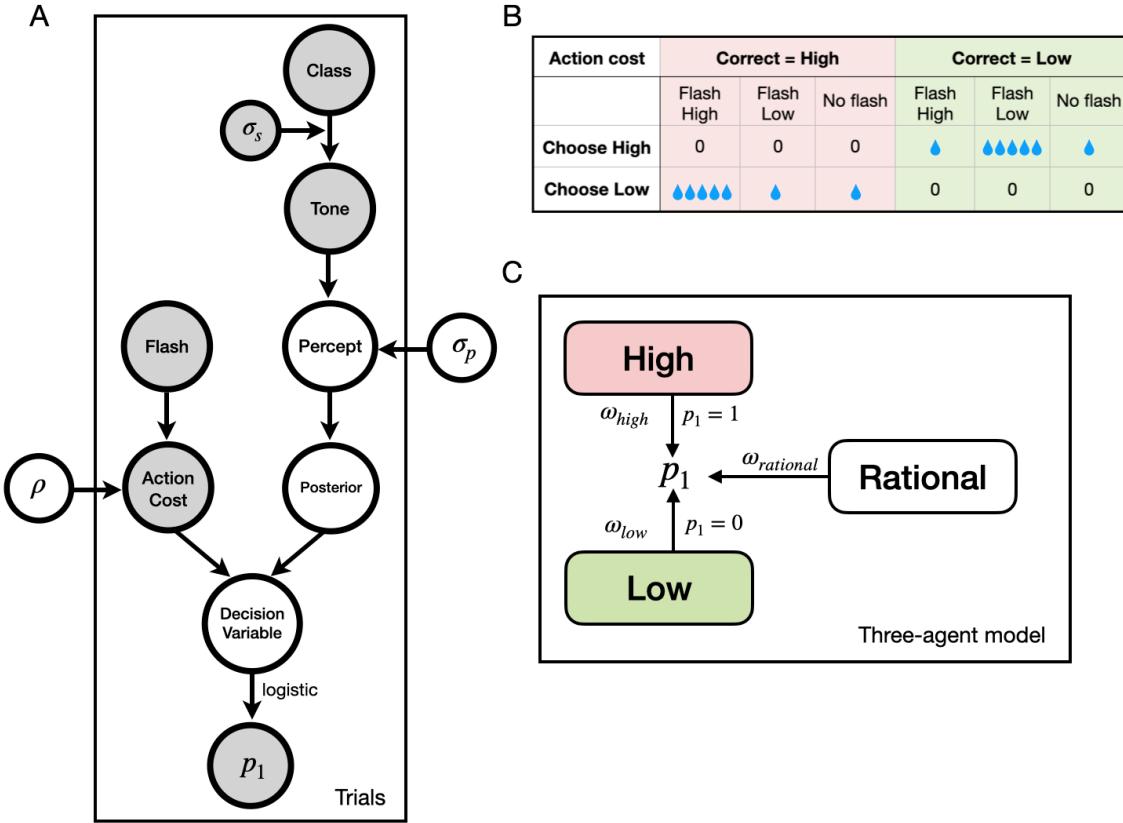
Next, we will briefly describe the model (FIGURE 4.4A, see modeling details in Methods). We start with the generative model, which specifies how the subject makes an observation  $x$  given the stimulus frequency  $s$  on each trial (FIGURE 4.1A). Recall that the stimulus was drawn from a truncated Gaussian centered at 3 with standard deviation  $\sigma_s$ , which was set by the experimenter. Bayesian models assume that through experience, subjects learn this distribution, and utilize it when inferring the correct class (low or high) given the tone on each trial. Thus, the observation distribution is dependent on two parameters:  $\sigma_s$ , which is known, and  $\sigma_p$  as in  $p(x|s) \sim \mathcal{N}(s, \sigma_p)$ , denoting the perceptual noise of each animal. The observation is then combined with class prior ( $p = 0.5$  for each class) to compute the class posterior, representing the animal's belief of each class given just the perceptual cue. To incorporate value information, we constructed a cost function where the choice is mapped to an action cost under different flash conditions (FIGURE 4.4B, TABLE 4.1). For example, when the high side is flashing and the correct class is high and the animal chooses low, the action cost would be  $base\ reward \times \kappa$ , a miss of considerable size. We included an additional parameter  $\rho$  as the exponent on the action cost, which is

equivalent to the curvature of the animal’s utility function:  $U = V^\rho$ , where  $U$  denotes utility and  $V$  is value. Finally, the  $\rho$ -adjusted action cost of choosing each class is integrated with its class posterior as the decision variable (EQUATION 4.15), which is transformed into a probability of choosing the high port using a *logistic* function.

We observed that some animals exhibited a constant, stimulus-independent rate of error known as ‘lapse’. Recently, it has been suggested to reflect exploration in a changing environment (Pisupati et al., 2019). To account for the lapses, we developed a ‘three-agent’ model that includes a ‘rational’ agent that outputs the probability of choosing the high port from the BDT model, a habitual ‘high’ agent that always chooses the high port, and a ‘low’ agent that always chooses the low port (FIGURE 4.4C). The choice on each trial is thus a weighted outcome of the votes from three agents with their respective mixing weights  $\omega$ , each implementing a different behavioral strategy. We refer to the final hybrid model as the ‘mixture-BDT’ model.

#### 4.3.4 The mixture-BDT model is insufficient to account for subjects’ behavior

We first validated that the model can correctly recover generative parameters from synthetic data (FIGURE 4.8). We estimated the joint posterior over the parameters for each animal separately using Hamiltonian Monte Carlo sampling in Stan (see example animals in FIGURE 4.5A, see all animals in FIGURE 4.10). Details of the modeling, including the priors, can be found in the Methods section. Overall, the animals all had a concave utility function ( $\rho = 0.30 [0.04 1.39]$ , median and 95% C.I. of concatenated posteriors across animals). They had medium to low levels of perceptual noise ( $\sigma_p = 0.25 [0.17 0.45]$ ), indicating that on average, they were sensitive

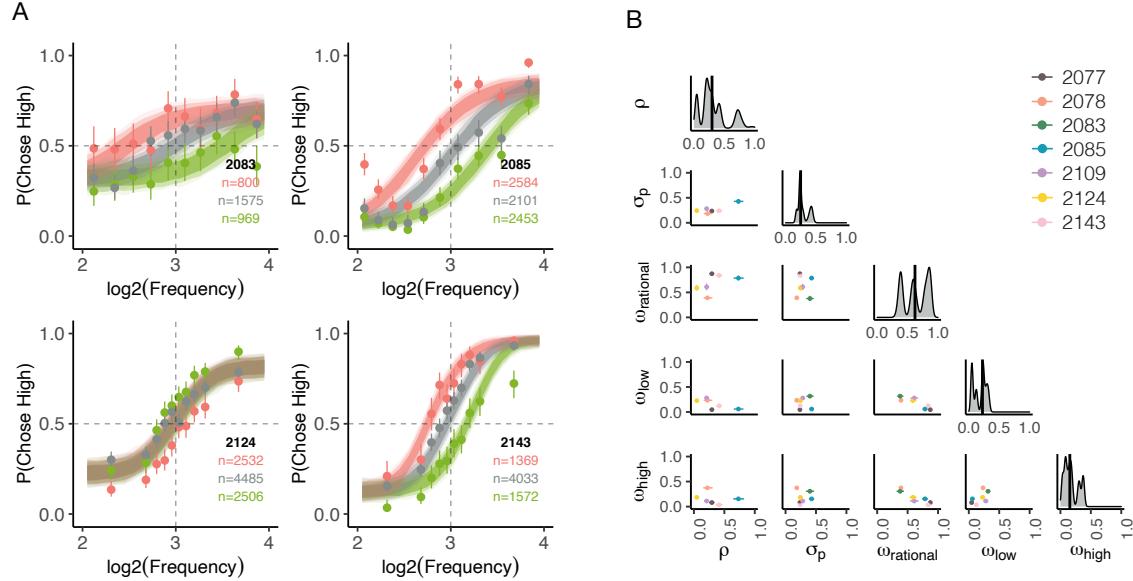


**Figure 4.4.** The mixture-BDT model. **A.** Graphical representation of just the BDT model. Using the plate notation, the variables shaded in gray are known or observable to the experimenters, and the variables in white are latent variables. The arrows indicate dependency between the variables. The two parameters to estimate are  $\rho$ , the curvature of utility function, and  $\sigma_p$ , the perceptual noise. See Methods for details. **B.** The cost function table. The cells where the correct choice is ‘high’ are shaded in red, and the correct choice is ‘low’ are shaded in green. Action cost is defined as the cost, or missed reward, from choosing the incorrect port. Naturally, it is 0 when the animal chooses correctly. It is the *base reward* (denoted by a single water drop) when the animal chooses incorrectly and the flash is not on the other side. It is *base reward*  $\times \kappa$  (denoted by multiple water drops) when the animal chooses incorrectly and the flash is on the other side. **C.** The three-agent mixture model. The animal’s final choice is modeled as a weighted average of the three agents, each implementing a different behavioral strategy to perform the task. The ‘rational’ agent outputs the probability of choosing high from the BDT model; the ‘high’ agent always chooses the high port ( $p_1 = 1$ ) and the ‘low’ agent always chooses the low port ( $p_1 = 0$ ). The final probability of choosing high port ( $p_1$ ) is a weighted sum with the agent’s respective weight  $\vec{\omega}$ , where  $\sum \vec{\omega} = 1$ .

to tone frequencies roughly 1.18 kHz apart. Consistent with GLMM results, animals with a sharper psychometric curve (e.g. 2143, pink dot in FIGURE 4.5B) had a smaller  $\sigma_p$  than animals with a flatter psychometric curve (e.g. 2083, green dot in FIGURE 4.5B). 2077, 2085 and 2143 were guided mostly by the rational agent ( $\omega_{rational} = 0.84$  [0.75 0.88],  $\omega_{low} = 0.06$  [0.04 0.14],  $\omega_{high} = 0.08$  [0.02 0.17], concatenated posteriors across these animals). In contrast, 2078, 2083, 2109 and 2124 displayed high levels of stimulus-independent bias ( $\omega_{rational} = 0.49$  [0.34 0.66],  $\omega_{low} = 0.25$  [0.21 0.34],  $\omega_{high} = 0.24$  [0.09 0.39]). However, this model failed to account for several aspect of animals' behavior. First, there is only one parameter,  $\rho$ , modulating how much subjects shift responses on PG trials, so the model predicts that the flash-induced shift should be symmetrical for left-flash and right-flash trials, which is not the case in our data. Second, the model predicts that flashes should result in horizontal shifts in the psychometric curve: the shift should depend on the perceptual uncertainty (FIGURE 4.11). In our data, some subjects shifted vertically (FIGURE 4.5A, 2124): a stimulus independent shift.

To quantify the degree to which the mixture-BDT model failed to fit the data, we refit the data, treating the perceptual trials, left flash and right flash trials as separate datasets, and fit each with a four-parameter sigmoid function (EQUATION 4.27, see details in Methods). If, as the mixture-BDT model predicts, flashes induces horizontal shifts (and small increases in slope), then the intercept term,  $x_0$ , would change the most, with small changes in the slope,  $b$ . However, in most animals,  $w_1$  or  $w_2$  changed in the flash trials relative to the perceptual trials, indicating vertical rather than horizontal shifts (FIGURE 4.6). Taken together, the modeling result suggests that the animal behavior is not well described by the normative BDT model, even

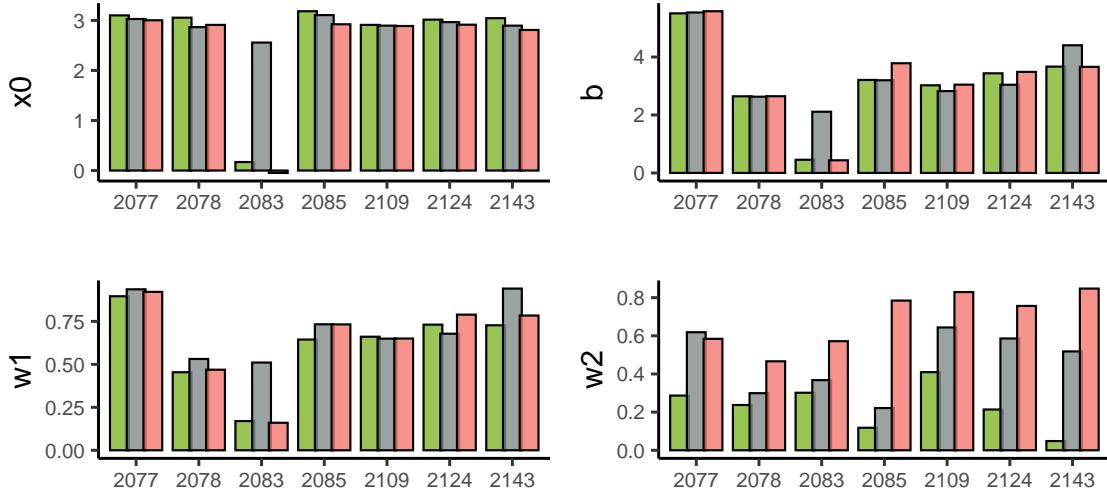
after taking into account lapse, and thus the animals were not optimally integrating the cues.



**Figure 4.5.** The mixture-BDT model fits. **A.** Example predictions using the model estimates. The circles with error bars are the binned mean  $\pm 95\%$  binomial confidence intervals from data. The colors represent the flashing condition, with red = high side flash, green = low side flash, gray = no flash. The ribbons are model predictions generated using the fitted parameters. The dark, medium and light shade represent  $\pm 80\%$ ,  $\pm 95\%$  and  $\pm 99\%$  confidence intervals, respectively. The model fit some animals fine (e.g. 2085, 2143) but was not able to fit other animals exhibiting vertical shifts due to flashes (e.g. 2124). **B.** Summary of the parameters of 7 animals. The mean and  $\pm 90\%$  confidence interval of each parameter pair are shown in the off-diagonal, colored by subject. Density plots of all fit posterior samples ( $n = 4000 \times 7$ ) for each parameter are on the diagonal, the black bar is the median.

#### 4.3.5 The animals are not reward-maximizing

The vertical shifts induced by the flashes are sub-optimal. How much reward were the subjects missing out on? To answer this, we compared animals' rewards obtained with an optimal, reward-maximizing Bayesian decision-maker. We define 'optimality'



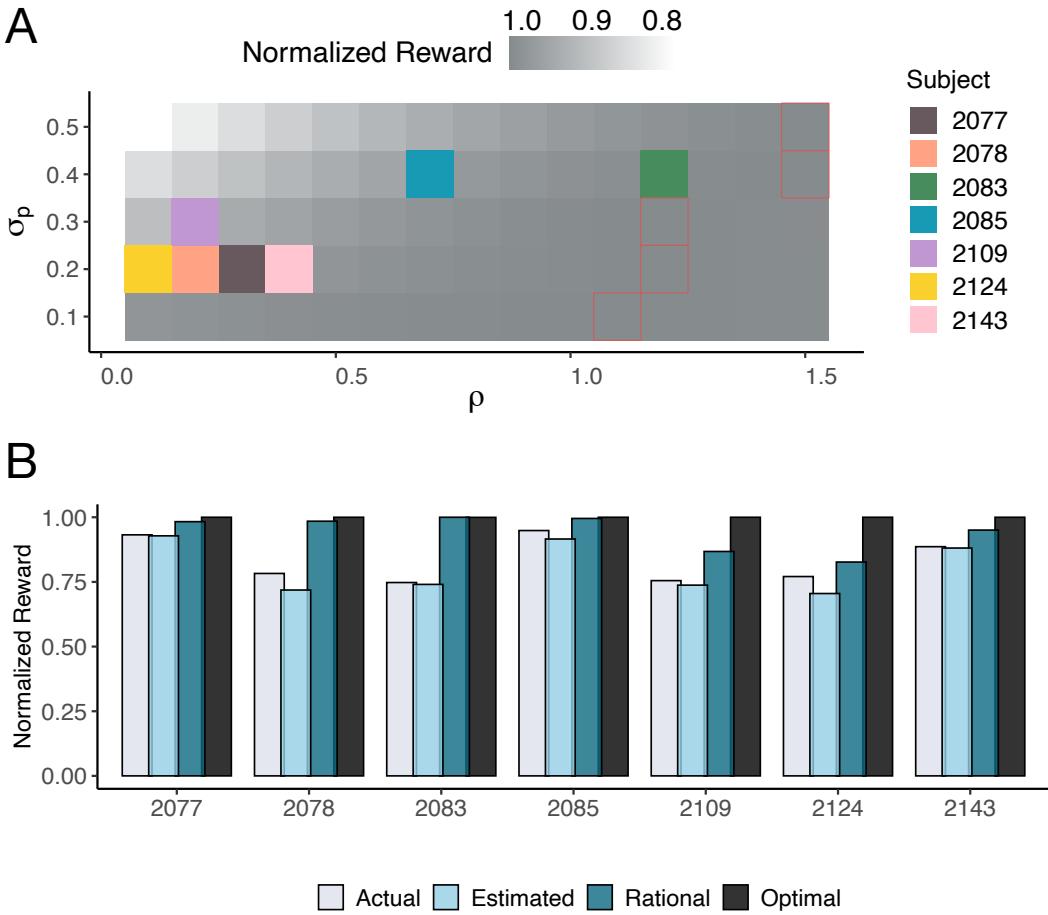
**Figure 4.6.** The four-parameter sigmoid model fits, shown in maximum a posteriori estimates. The model was fit separately on data from each condition: red - high side flash, green - low side flash, gray - no flash.

here as obtaining the maximum possible reward given a fixed perceptual noise  $\sigma_p$  in a particular task dataset. Interestingly, it was found that in order for a purely rational BDT agent to obtain maximum reward, its utility curvature  $\rho$  needs to balance with its perceptual noise  $\sigma_p$  (FIGURE 4.7A). This can be intuitively understood by going through some examples. For a subject with large  $\sigma_p$ , it will have a high error rate due to poor perceptual judgment, the strategy to maximize reward would be to choose the flashing side as much as possible and result in a convex utility function. Alternatively, a subject with very small  $\sigma_p$  will get most trials correct anyway, it does not need to ‘value’ the flash more than what it represents, resulting in a close-to-linear utility function. Another interesting result from the simulation analysis is that  $\rho$  does not affect the total reward much when  $\sigma_p$  is small, but its value plays a big role when  $\sigma_p$  is large. This in part, explains why subjects like 2077 and 2143 are closest to

the ‘optimal’, reward-maximizing agent even with  $\rho$  smaller than the best  $\rho$  (FIGURE 4.7B). For subjects 2078, 2109 and 2124, their large stimulus-independent bias seemed to be the culprit for obtaining lower reward overall. A fascinating example is 2083, its estimated  $\rho$  and  $\sigma_p$  combination are close to being optimal (FIGURE 4.7A, GREEN SQUARE). The fact that it only obtained 75% of the maximum reward was entirely due to its lapse rate, as a rational agent with its  $\rho$  and  $\sigma_p$  obtained just as much reward as the optimal agent. On average, the animals obtained  $83.2 \pm 3.3\%$  of the maximum reward obtained by their respective optimal agent, modeled with the same perceptual noise. Taken together, the analysis showed that the animals are not optimal in reward-maximization in the task, and this is due to a combination of high lapse rate and extreme risk-aversion.

## 4.4 Discussion

Decision-making is a term referring to the integration and transformation of external information with internal beliefs into an action. The external information may contain perceptual as well as value aspects of the decision required at hand. Despite its importance, only few studies have examined the behavioral and neurobiological underpinnings of decisions that involve percept-value integration. Here, we developed the perceptual gambling task where the rat made choices informed by both perceptual (tone frequency) and value (light flash) cues, on a trial-by-trial basis. Although the subjects did not, on average, shift their choices symmetrically by value as in monkeys (Rorie et al., 2010), the animals nevertheless showed sensitivity to flashes. We characterized behavior using the Bayesian decision theory, which assumes an opti-



**Figure 4.7.** The animals were sub-optimal in obtaining reward. **A.** Different values of  $\rho$  is required for different  $\sigma_p$  to obtain maximum reward. For each  $\rho$  and  $\sigma_p$  combination, we simulated a rational BDT agent ( $\vec{\omega} = [1, 0, 0]$ ) and computed the reward it obtained using the same task dataset. The color of the heatmap corresponds to the total reward obtained by this parameter combination, normalized across rows. The MAP estimates of  $\rho$  and  $\sigma_p$  for each subject with the three-agent BDT model are shown in colored blocks. The  $\rho$  that resulted in maximum reward given each  $\sigma_p$  is highlighted in red squares. **B.** Comparing reward obtained using four different agents. Actual: the sum of actual reward obtained for this animal. Estimated: the total reward obtained by a BDT agent simulated with fitted  $\rho$ ,  $\sigma_p$ ,  $\vec{\omega}$ . Rational: the total reward obtained by a BDT agent simulated with fitted  $\rho$ ,  $\sigma_p$  and  $\vec{\omega} = [1, 0, 0]$ . Optimal: the total reward obtained by a BDT agent simulated with fitted  $\sigma_p$ , the best  $\rho$  from **A** given  $\sigma_p$ , and  $\vec{\omega} = [1, 0, 0]$ . All normalized by the maximum reward, which was obtained by the optimal agent. See details in Methods.

mal integration of individual perceptual noise and reward sensitivity, as well as the statistics of the task environment. It was found that the behavior was not well fit by the BDT model, even after accounting for lapses, because subjects responded to the flashes by shifting their choices in a stimulus independent way. Finally, we quantified the fraction of the reward that animals were foregoing by using a sub-optimal strategy. Model-based analysis revealed that the missing reward was due to a large lapse rates and risk aversion (extremely concave utility functions).

Overall, the results show that the animals are not behaving optimally in the task. Their choices were influenced by the previous trial's outcome, even while the tone frequency and flash was independent across trials. Lak et al. (2020) also observed strong history effects, but in that task the animal was encouraged to incorporate history information due to the block design. Although Bayesian decision theory has had success in explaining and predicting human behavior (e.g. Cogley and Sargent, 2008, Körding and Wolpert, 2006), but it has rarely been used to quantify rodent behavior. One key assumption of Bayesian decision models is that the subject has *learned* the distributions of latent variables in the task environment and utilized them fully when making inference. It is likely that both the robust history effect and poor fit of the mixture-BDT are a result of incomplete or ongoing learning. For most animals, the training process involved adjusting perceptual difficulty  $\sigma_s$  and reward multiplier  $\kappa$  periodically to induce significant shifting. As a result, the animals may have internalized the environmental volatility and were actively exploring reward contingencies. This may also underlie the substantial lapse rate observed in some animals (Pisupati et al., 2019). Thus, we emphasize that the poor performance of BDT does not suggest rodents are non-Bayesian agents, it may merely reflect the lack

of learning aspects in the theory. In fact, it seems that the monkey performance in a similar task from Rorie et al. (2010) can be well fit by the BDT, as their value-induced shift in the psychometric curves was horizontal.

The main takeaway from the training process was that task parameters like perceptual difficulty and reward multiplier heavily interacted with the animal’s sensory noise  $\sigma_p$  and utility exponent  $\rho$ . Future researchers interested in adopting this framework are encouraged to use a model-based training method. Specifically, when  $\sigma_p$  is low and  $\rho$  is small, to induce a behavioral shift, the experimenter needs to increase perceptual difficulty and increase the reward multiplier. When  $\sigma_p$  is high and  $\rho$  is large, to prevent the animal from simply choosing the side with flash and ignoring the sound, the experimenter can reduce perceptual difficulty and decrease the reward multiplier. The animal’s  $\sigma_p$  can be estimated from the performance on the perceptual trials alone, prior to any value training. Estimating  $\rho$  is challenging without training the animal on a task that exposes its subjective utility function. Nonetheless, results from choice under risk showed that most rats have concave utility functions ( $\rho < 1$  Zhu et al., 2021, Constantinople et al., 2019b). On that account, it is reasonable to assume a small  $\rho$  in the beginning of training unless the animal showed otherwise. However, we do not exclude the possibility that the animal may ‘adapt’ the shape of its utility function in different contexts, for example, a noisy perceptual decision-maker may deliberately become more risk-seeking to harvest more reward. There is some mixed evidence from behavioral ecology experiments supporting a context-dependent change in utility concavity (Kacelnik and El Mouden, 2013). Finally, given the small  $\sigma_p$  found in most animals in this study, it is advisable to use a narrower range of auditory stimuli (e.g. 5.65 - 11.31 kHz) to facilitate the training process.

In this manuscript, we did not explicitly test whether the flash-induced shift was due to a perceptual bias or a response bias, which predicts that the value information exerts influence *on* or *separate from* sensory processing, respectively. Nonetheless, the perceptual gambling task along with the mixture-BDT model together, can generate specific hypotheses on how to distinguish these two scenarios. Future researchers can record activity from the secondary motor regions and associative sensory regions in well-trained rats to establish correlative relationships. A response bias would predict increased activity in the motor region when the chosen side is cued for higher reward, whereas a perceptual bias would predict differential activity elicited by the same auditory stimulus in the sensory areas under flash and no-flash conditions. Furthermore, causal evidence for either scenario can be obtained with pharmacological and optogenetic inactivations. If by inactivating the secondary motor region the animal simply shifts less to the flashing side, which is equivalent to a decreased utility exponent  $\rho$  in the model, the response bias hypothesis will be supported. Even more interestingly, this would suggest a dissociable process of perceptual decision and value computation. On the other hand, the perceptual bias will be supported if the animal shifts less by following inactivation of its associative sensory areas.

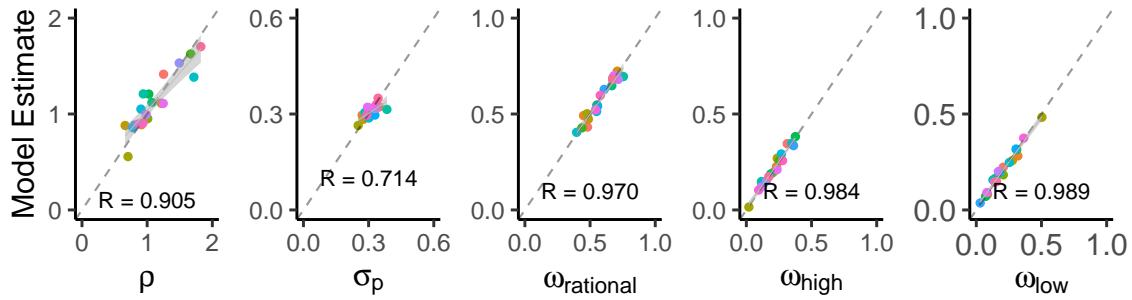
Another promising avenue of research is to use the perceptual gambling task to study confidence. Confidence is generally defined as the degree of belief in the truth of a proposition or the reliability of a piece of information, be it memory, observation or decision (Kepecs and Mainen, 2012). One important nuance is that there confidence has to be about a specific belief. In the PG task, one can distinguish perceptual confidence from decision confidence. In signal detection theory, perceptual confidence is defined as the distance from an internal representation of a stimulus to the decision

boundary. On average, accuracy is a proxy for perceptual confidence (Clarke et al., 1959, Galvin et al., 2003). This is a model of perceptual confidence, but it becomes indistinguishable from decision confidence in simple perceptual decision tasks. Recently, a Bayesian equivalent was proposed under the Bayesian decision theory, which defined confidence as the observer's posterior probability of being correct (Kepcs and Mainen, 2012, Hangya et al., 2016, Pouget et al., 2016). Using the Bayesian definition of confidence, research groups were able to find behavioral and neural correlates that resemble the qualitative signatures of Bayesian confidence in rats (Kepcs et al., 2008, Lak et al., 2014) and humans (Sanders et al., 2016). We observe that the Bayesian decision theory framework can distinguish the two, such that the perceptual confidence is simply the log ratio of posterior *sans* action cost, and the decision confidence is the log ratio of posterior with action cost. Our task makes investigating the neural underpinnings of these two kinds of confidence possible. Specifically, areas computing perceptual confidence should not vary across flashing conditions, as the perceptual boundary remains the same (unless, of course, the change in value causes a change in percept). In contrast, areas responsible for decision confidence should show differential activities across conditions, as the asymmetric reward would shift the indifference point, entailing a considerable change in decision confidence even with the same perceptual stimulus. Activities in these areas can be related to the Bayesian confidence measures as a test of theory, although it has been cautioned that the use of qualitative match is problematic (Adler and Ma, 2018).

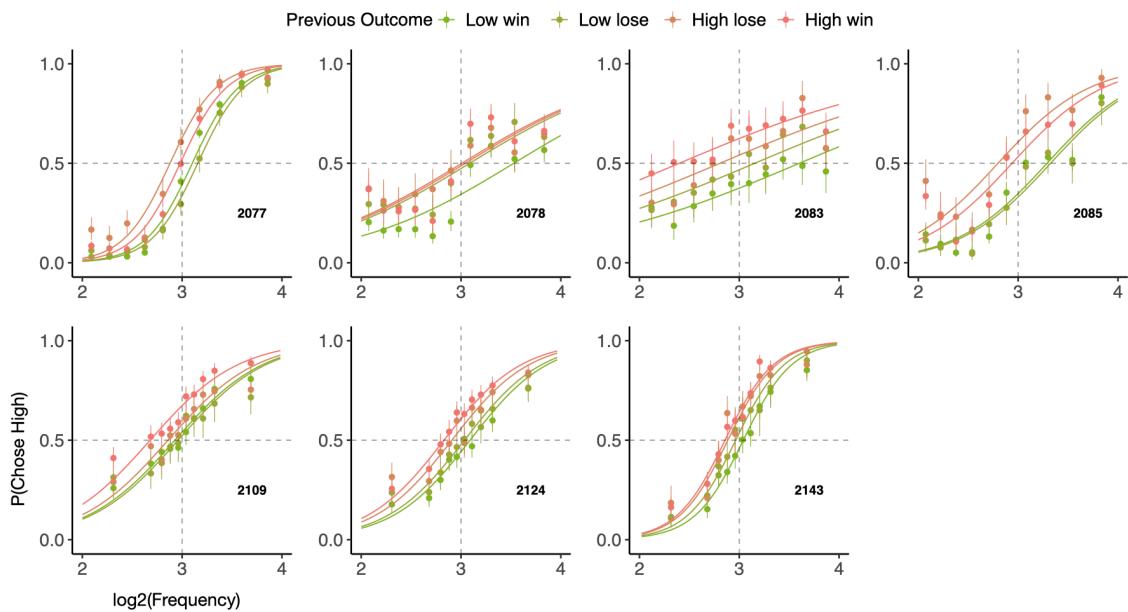
In conclusion, we present the perceptual gambling task as a proof of concept, demonstrating that a integration of perceptual and value cues on a trial-by-trial basis is possible for rats. Future researchers interested in percept-value integration and

confidence are encouraged to adopt this framework. Using model-based analysis, the brain circuits underlying these behavior can be rigorously explored with testable hypotheses.

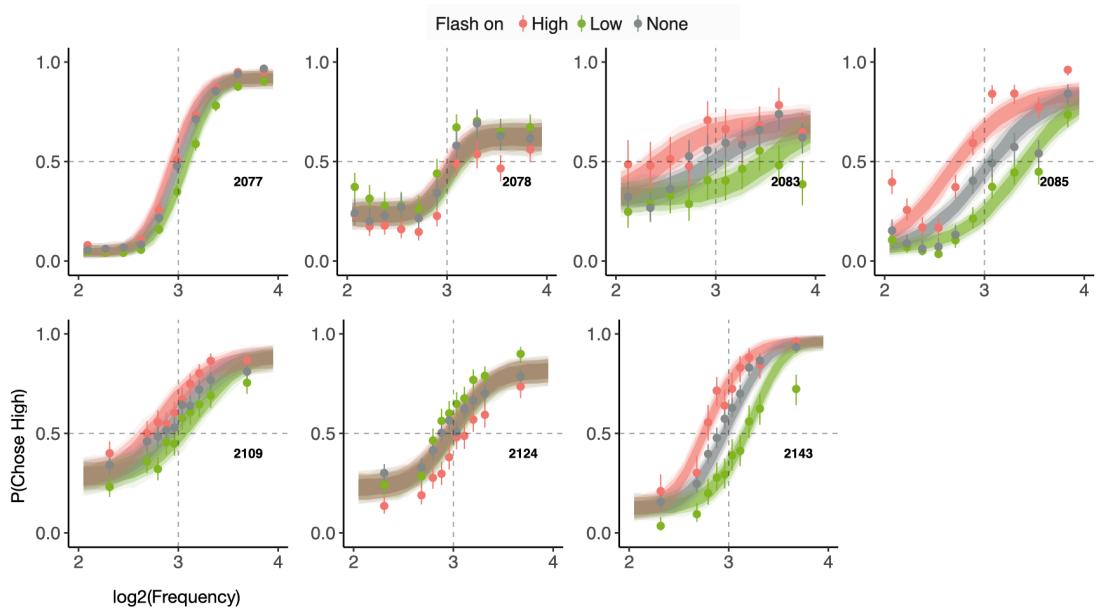
## 4.5 Supplementary figures



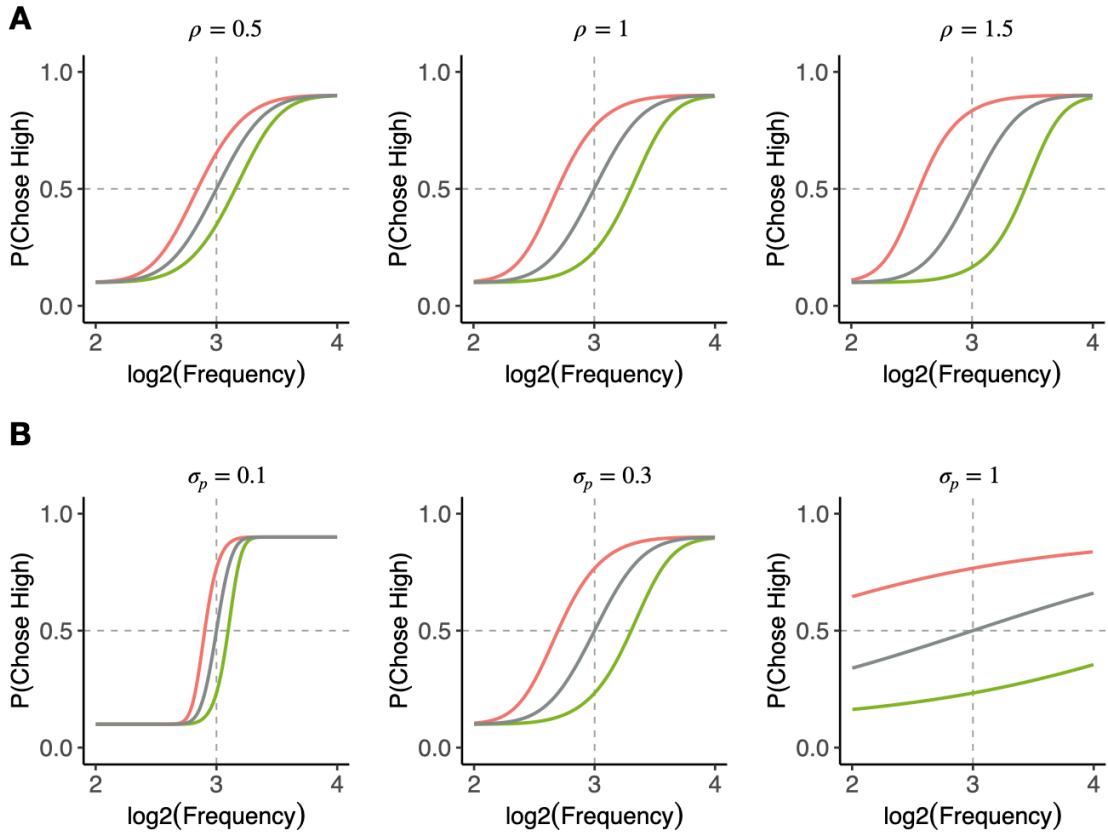
**Figure 4.8.** The mixture-BDT model can recover data-generating parameters accurately. Twenty Synthetic datasets were created by sampling from the same prior distributions as specified in Methods. The true parameter value is on the x-axis, the maximum *a posteriori* estimation is on the y-axis. Color represents the identity of each synthetic dataset. All the parameters fall along the diagonal line (all  $R > 0.7, p < 0.001$ , Pearson's correlation test).



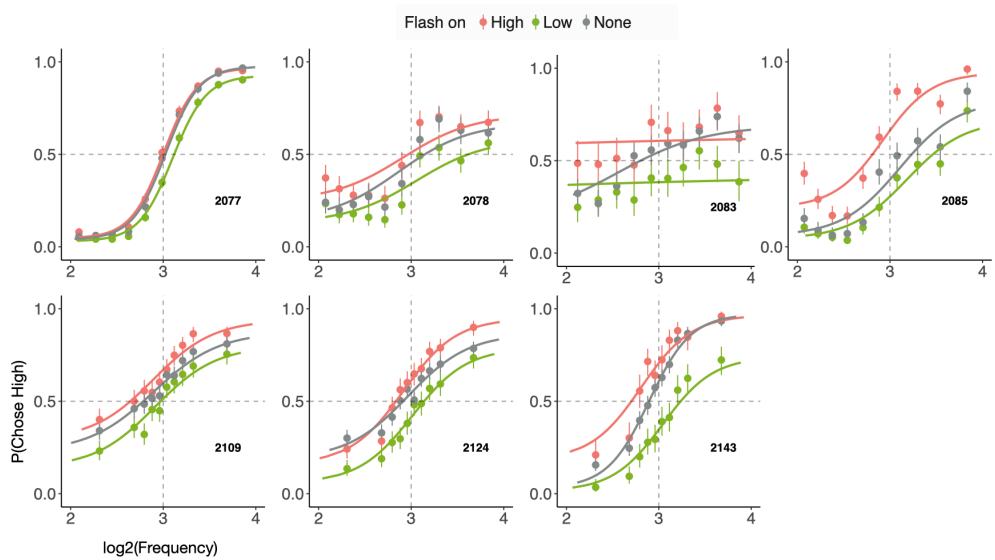
**Figure 4.9.** Animal's choices were influenced by the outcome from the previous trial. The circles with error bars are the binned mean and 95% binomial confidence intervals. The lines are generated by a generalized linear model. The colors represent the outcome of the previous trial.



**Figure 4.10.** The mixture-BDT model fits for each subject. The circles with error bars are the binned mean and 95% binomial confidence intervals. The lines are generated from a synthetic sigmoid agent using maximum *a posteriori* parameter estimates. The colors represent the flashing condition, with red = high side flash, green = low side flash, gray = no flash.



**Figure 4.11.** The mixture-BDT model only allows for symmetric and horizontal shifts as a result of asymmetric reward. The data is simulated from a BDT agent with  $\rho = 1$ ,  $\sigma_p = 0.3$ ,  $\omega_{rational} = 0.8$ ,  $\omega_{high} = 0.1$ ,  $\omega_{low} = 0.1$  unless otherwise specified. The colors represent the flashing condition, with red = high side flash, green = low side flash, gray = no flash. **A.** Different levels of  $\rho$ , the utility exponent, determine the amount of horizontal shift without changing the slope. **B.** Different levels of  $\sigma_p$ , the perceptual noise, change both the amount of horizontal shift and slope of the psychometric functions.



**Figure 4.12.** The four-parameter sigmoid model fits for each subject. The circles with error bars are the binned mean and 95% binomial confidence intervals. The lines are generated from a synthetic sigmoid agent using maximum *a posteriori* parameter estimates. The colors represent the flashing condition, with red = high side flash, green = low side flash, gray = no flash.

# Chapter 5

## Conclusion

### 5.1 General summary

In this dissertation, I have presented results from three studies, all of which contributed to the understanding of decisions under risk in rats. CHAPTER 2 AN ANALYSIS OF DECISION UNDER RISK IN MICE, RATS AND HUMANS established a cross-species paradigm for risky choice. Mice, rats and human subjects chose between a safe and a lottery option, whose magnitude (and probability) was signaled by a distinct auditory cue on a trial-by-trial basis. By systematically modeling choices with several functional forms of risk, noise specification, heuristic mixture and history parameters, my modeling results highlighted a crucial distinction in the strategies used by rodents and humans. 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 I isolated the history-dependence and stimulus-independent biases, the distribution of utility func-

tions and variance aversion across species became more overlapped, supporting the use of rodents to examine the neurobiology of risk-tolerance. These results provide not only a first qualification of variance-aversion in mice and rats on a population scale, but also a first direct cross-species comparison of risk attitude known to date.

In **CHAPTER 3 FRONTAL BUT NOT PARIETAL CORTEX IS REQUIRED FOR DECISIONS UNDER RISK**, I tested the causal contributions of two cortical areas, the FOF and PPC, in risky choice. While both regions were previously implicated in decision-making, pharmacological inactivations produced strikingly different results in rats with stable behavior. As in perceptual decisions, silencing the PPC produced minimal effects on risky choice, even while performance on an interleaved free-choice task was impaired. In contrast, inactivating the FOF shifted the choices substantially to the fixed surebet. The change was best characterized by a decrease in  $\rho$  and not by any heuristic mixture parameters, suggesting the animal's utility computation was affected by FOF inactivation. This was parsimoniously explained by a dynamical model where the FOF is part of network that performs sensory-to-value transformations. These results not only represent the first evidence against a causal role of PPC in economic choices, but also highlight the role of premotor cortex and support a distributed circuit for economic choices.

In **CHAPTER 4 A RODENT PARADIGM FOR STUDYING PERCEPTUAL DECISIONS UNDER ASYMMETRIC REWARD**, I presented rat behavior in a task that required the integration of both perceptual and value cues on a trial-by-trial basis. The perceptual gambling task, as it was named, effectively invited the animal to 'gamble' its confidence of the perceptual stimulus for a higher reward. Risk was conveyed in the task as each stimulus, after repetitive training, corresponded to perceptual confidence,

which should be accessible to the animals. It was found that despite sensitivity to the perceptual and value cues, their behavior was not optimal: subjects tended to shift their choices in a stimulus-independent way following value cues. Moreover, subjects tended to under-shift, which could be interpreted as being over-confident in their perceptual beliefs or as being extremely risk-averse. Results from this chapter showcased that an integration of perceptual and value cues on a trial-by-trial basis is possible for rats, and suggested several exciting research topics enabled by the perceptual gambling task.

## 5.2 Limitations and recommendations

Although the risky choice task varied lottery magnitude independently across trials, it only had a fixed lottery probability. The initial design was to include various probabilities signaled by sound amplitude as in the human task, however, it proved difficult to train the animals (especially mice) on this version. Results from Constantinople et al. (2019b) demonstrated that rats can choose under varying lottery magnitudes and probabilities, where the magnitude was cued by auditory clicks and the probability by light flashes. In retrospect, we should have used multisensory (auditory + visual) rather than just auditory cues to convey the information.

In CHAPTER 3, I used pharmacological inactivations to test the causal role of the FOF and PPC. Due to its tendency to spread, I cannot rule out the possibility that observed impairments from FOF inactivation was partially due to muscimol spillover into adjacent regions, especially the anterior Cingulate Cortex (ACC). Neurons in monkey ACC are shown to encode and integrate lottery magnitude and probability

(Hayden and Platt, 2010, Kennerley et al., 2009). In a matching pennies task where rats played against a virtual competitor, disturbing the locus coeruleus input into ACC made their behavior more stochastic than strategic (Tervo et al., 2014). In general, there is consensus that ACC is a key region for adaptive action selection and decision making (Monosov et al., 2020). Although evidence suggest that silencing ACC will likely produce similar effects reported in CHAPTER 3, I argue that this was not the case. First, inactivation of rat ACC did not impair their behavior in a rodent Iowa gambling task, where the optimal strategy is to integrate long-term probability and outcome associated with each port (Zeeb et al., 2015). Second, histology showed that the actual FOF targets were closer to M1 than ACC in all 8 rats (FIGURE 3.10). Lastly, preliminary data using optogenetic inactivation in the FOF showed a similar shift towards choosing the surebet in rats, albeit to a smaller degree (data collected by Chaofei Bao). Overall, the conclusion remains that FOF is a key node in a network that performs sensory-to-value transformations.

The perceptual gambling task described in CHAPTER 4 was originally proposed to be the main paradigm with which I would investigate the neural mechanisms of percept-value integration, a process that we know very little of. However, the task turned out to be extremely parameter-sensitive and difficult to train the animals with, hence I only collected some not-so-ideal behavior. In retrospect, of course it was difficult to train that way. The task required the animal to trade off a delicate balance between its perceptual noise and reward sensitivity, both of which the experimenter knew nothing of until a behavioral shift was observed. An impossible loop! By all means, I do not regret ‘wasting’ quite some time on getting this behavior to work. Designing a behavior of choice is not easy, it requires not only knowledge of the

subject's inner states: motivation, reward sensitivity, perceptual abilities etc. and how they interact, but also an ability to take the perspective of a thirsty rat in the chamber. This experience helped me become a better behavioral scientist, and grew genuine appreciation of people who successfully developed highly complex rodent behavior to tackle interesting problems (e.g. Miller et al., 2017, Constantinople et al., 2019a, Duan et al., 2015, Kepecs et al., 2008).

Future researchers interested in adopting this task can improve the training process in several ways. First, the animal will more likely shift to the flashing side when its perceptual confidence is low. Regardless of the animal's perceptual noise, I recommend starting with a narrower frequency range than 4 - 16 kHz, then extend or contract the range based on performance in perceptual trials. Second, once the animal reliably differentiates the tones, its perceptual noise can be estimated by a simple Bayesian model. By plugging the perceptual noise term in the BDT model, the future researcher will find a minimum threshold of the flash reward multiplier that will fall within the 'sweet spot' to induce a perceptual gambling effect (assuming  $\rho < 1$ ). Lastly, other perceptual stimuli can be used instead of tone frequency, such as the number of auditory clicks (as in Erlich et al., 2015) and the contrast of visual gratings (as in Lak et al., 2020).

# Bibliography

23and Me Research Team, eQTLgen Consortium, International Cannabis Consortium, Social Science Genetic Association Consortium, R. Karlsson Linnér, P. Biroli, E. Kong, S. F. W. Meddends, R. Wedow, M. A. Fontana, M. Lebreton, S. P. Tino, A. Abdellaoui, A. R. Hammerschlag, M. G. Nivard, A. Okbay, C. A. Rietveld, P. N. Timshel, M. Trzaskowski, R. de Vlaming, C. L. Zünd, Y. Bao, L. Buzdugan, A. H. Caplin, C.-Y. Chen, P. Eibich, P. Fontanillas, J. R. Gonzalez, P. K. Joshi, V. Karhunen, A. Kleinman, R. Z. Levin, C. M. Lill, G. A. Meddends, G. Muntané, S. Sanchez-Roige, F. J. van Rooij, E. Taskesen, Y. Wu, F. Zhang, A. Auton, J. D. Boardman, D. W. Clark, A. Conlin, C. C. Dolan, U. Fischbacher, P. J. F. Groenen, K. M. Harris, G. Hasler, A. Hofman, M. A. Ikram, S. Jain, R. Karlsson, R. C. Kessler, M. Kooyman, J. MacKillop, M. Männikkö, C. Morcillo-Suarez, M. B. McQueen, K. M. Schmidt, M. C. Smart, M. Sutter, A. R. Thurik, A. G. Uitterlinden, J. White, H. de Wit, J. Yang, L. Bertram, D. I. Boomsma, T. Esko, E. Fehr, D. A. Hinds, M. Johannesson, M. Kumari, D. Laibson, P. K. E. Magnusson, M. N. Meyer, A. Navarro, A. A. Palmer, T. H. Pers, D. Posthuma, D. Schunk, M. B. Stein, R. Svento, H. Tiemeier, P. R. H. J. Timmers, P. Turley, R. J. Ursano, G. G. Wagner, J. F. Wilson, J. Gratten, J. J. Lee, D. Cesarini, D. J. Benjamin, P. D.

Koellinger, and J. 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, Feb. 2019. ISSN 1061-4036, 1546-1718. doi: 10.1038/s41588-018-0309-3.

A. Abrahamyan, L. L. Silva, S. C. Dakin, M. Carandini, and J. 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.

W. T. Adler and W. J. Ma. Limitations of proposed signatures of Bayesian confidence. Jan. 2018. doi: 10.1101/218222.

V. Aguijón-Rodríguez, D. Angelaki, H. Bayer, N. Bonacchi, Carandini.M, F. Cazettes, G. Chapuis, A. K. Churchland, Y. Dan, E. Dewitt, M. Faulkner, H. Forrest, L. Haetzel, M. Hüsser, S. B. Hofer, F. Hu, A. Khanal, C. Krasniak, I. Laranjeira, Z. F. Mainen, G. Meijer, N. J. Miska, T. D. Mrsic-Flogel, M. Murakami, J. P. Noel, A. Pan-Vazquez, C. Rossant, J. Sanders, K. Socha, R. Terry, A. E. Urai, H. Vergara, M. Wells, C. J. Wilson, I. B. Witten, L. B. Wool, and A. M. Zador. The International Brain Laboratory. *eLife*, 10, May 2021. doi: 10.7554/eLife.63711.

S. H. Ahmed. Individual decision-making in the causal pathway to addiction: Contributions and limitations of rodent models. *Pharmacology Biochemistry and Behavior*, 164:22–31, Jan. 2018. ISSN 0091-3057. doi: 10.1016/j.pbb.2017.07.005.

A. Akrami, C. D. Kopec, M. E. Diamond, and C. D. Brody. Posterior parietal cortex

represents sensory history and mediates its effects on behaviour. *Nature*, 554(7692):368–372, Feb. 2018. ISSN 1476-4687. doi: 10.1038/nature25510.

M. Allais. Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine. *Econometrica*, 21(4):503–546, 1953. ISSN 0012-9682. doi: 10.2307/1907921.

C. Allefeld, C. S. Soon, C. Bogler, J. Heinzle, and J.-D. Haynes. Sequential dependencies between trials in free choice tasks. *arXiv:1311.0753 [q-bio]*, Nov. 2013.

A. P. Anokhin, S. Golosheykin, J. Grant, and A. C. Heath. Heritability of Risk-Taking in Adolescence: A Longitudinal Twin Study. *Twin Research and Human Genetics*, 12(4):366–371, Aug. 2009. ISSN 1832-4274, 1839-2628. doi: 10.1375/twin.12.4.366.

K. J. Arrow. *Aspects of the Theory of Risk-Bearing*. Yrjö Jahnssonin Säätiö, 1965.

D. R. Bach and R. J. Dolan. Knowing how much you don’t know: A neural organization of uncertainty estimates. *Nature Reviews Neuroscience*, 13(8):572–586, Aug. 2012. ISSN 1471-0048. doi: 10.1038/nrn3289.

F. Balci, D. Freestone, and C. 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, Feb. 2009. ISSN 1091-6490. doi: 10.1073/pnas.0812709106.

N. Barberis and M. Huang. Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review*, 98(5):2066–2100, Dec. 2008. ISSN 0002-8282. doi: 10.1257/aer.98.5.2066.

- F. Barthaš and A. C. Kwan. Secondary Motor Cortex: Where ‘Sensory’ Meets ‘Motor’ in the Rodent Frontal Cortex. *Trends in Neurosciences*, 40(3):181–193, Mar. 2017. ISSN 0166-2236. doi: 10.1016/j.tins.2016.11.006.
- V. Bartoš, M. Bauer, J. Chytilová, and I. Levely. Psychological Effects of Poverty on Time Preferences. *The Economic Journal*, 131(638):2357–2382, Aug. 2021. ISSN 0013-0133, 1468-0297. doi: 10.1093/ej/ueab007.
- D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1):1–48, Oct. 2015. ISSN 1548-7660. doi: 10.18637/jss.v067.i01.
- T. E. J. Behrens, M. W. Woolrich, M. E. Walton, and M. F. S. Rushworth. Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9):1214–1221, Sept. 2007. ISSN 1097-6256. doi: 10.1038/nn1954.
- D. A. Bercovici, O. Princz-Lebel, M. T. Tse, D. E. Moorman, and S. B. Floresco. Optogenetic Dissection of Temporal Dynamics of Amygdala-Striatal Interplay during Risk/Reward Decision Making. *eNeuro*, 5(6), Dec. 2018. ISSN 2373-2822. doi: 10.1523/ENEURO.0422-18.2018.
- K. C. Berridge and J. P. O’Doherty. Chapter 18 - From Experienced Utility to Decision Utility. In P. W. Glimcher and E. Fehr, editors, *Neuroeconomics (Second Edition)*, pages 335–351. Academic Press, San Diego, Jan. 2014. ISBN 978-0-12-416008-8. doi: 10.1016/B978-0-12-416008-8.00018-8.
- J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A Fresh Approach

- to Numerical Computing. *SIAM Review*, 59(1):65–98, Jan. 2017. ISSN 0036-1445, 1095-7200. doi: 10.1137/141000671.
- S. Bhatia and G. Loomes. Noisy preferences in risky choice: A cautionary note. *Psychological Review*, 124(5):678, 2017. ISSN 1939-1471. doi: 10.1037/rev0000073.
- E. Brandstätter, G. Gigerenzer, and R. 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.
- B. W. Brunton, M. M. Botvinick, and C. D. Brody. Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science*, 340(6128):95–98, Apr. 2013. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1233912.
- D. J. Bucci. Posterior parietal cortex: An interface between attention and learning? *Neurobiology of Learning and Memory*, 91(2):114–120, Feb. 2009. ISSN 1074-7427. doi: 10.1016/j.nlm.2008.07.004.
- D. J. Bucci and A. C. Chess. Specific changes in conditioned responding following neurotoxic damage to the posterior parietal cortex. *Behavioral Neuroscience*, 119(6):1580, 2005. ISSN 1939-0084. doi: 10.1037/0735-7044.119.6.1580.
- Y. Burak and I. R. Fiete. Fundamental limits on persistent activity in networks of noisy neurons. *Proceedings of the National Academy of Sciences*, 109(43):17645–17650, Oct. 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1117386109.
- C. Burke and P. Tobler. Coding of Reward Probability and Risk by Single Neurons in Animals. *Frontiers in Neuroscience*, 5:121, 2011. ISSN 1662-453X. doi: 10.3389/fnins.2011.00121.

- X. Cai and C. Padoa-Schioppa. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *The Journal of Neuroscience*, 32(11):3791–3808, Mar. 2012. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.3864-11.2012.
- C. F. Camerer. An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2(1):61–104, Apr. 1989. ISSN 1573-0476. doi: 10.1007/BF00055711.
- M. Carandini and A. K. Churchland. Probing perceptual decisions in rodents. *Nature Neuroscience*, 16(7):824–831, July 2013. ISSN 1097-6256. doi: 10.1038/nrn.3410.
- B. Carpenter, A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 20:1–37, 2016.
- L. S. Carvalho, S. Meier, and S. W. Wang. Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday. *American Economic Review*, 106(2):260–284, Feb. 2016. ISSN 0002-8282. doi: 10.1257/aer.20140481.
- L. L. Chen and S. P. Wise. Supplementary eye field contrasted with the frontal eye field during acquisition of conditional oculomotor associations. *Journal of Neurophysiology*, 73(3):1122–1134, Mar. 1995. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.1995.73.3.1122.
- X. Chen and V. Stuphorn. Sequential selection of economic good and action in medial frontal cortex of macaques during value-based decisions. *eLife*, 4, Nov. 2015. ISSN 2050-084X. doi: 10.7554/eLife.09418.

- X. Chen and V. Stuphorn. Inactivation of Medial Frontal Cortex Changes Risk Preference. *Current Biology*, 28(19):3114–3122.e4, Oct. 2018. ISSN 0960-9822. doi: 10.1016/j.cub.2018.07.043.
- X. Chen, M. Zirnsak, G. M. Vega, and T. Moore. Frontal eye field neurons selectively signal the reward value of prior actions. *Progress in Neurobiology*, 195:101881, Dec. 2020. ISSN 0301-0082. doi: 10.1016/j.pneurobio.2020.101881.
- V. S. Chib, A. Rangel, S. Shimojo, and J. 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, Sept. 2009. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2575-09.2009.
- G. I. Christopoulos, P. N. Tobler, P. Bossaerts, R. J. Dolan, and W. Schultz. Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *Journal of Neuroscience*, 29(40):12574–12583, Oct. 2009. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2614-09.2009.
- F. R. Clarke, T. G. Birdsall, and W. P. Tanner. Two Types of ROC Curves and Definitions of Parameters. *The Journal of the Acoustical Society of America*, 31(5):629–630, May 1959. ISSN 0001-4966. doi: 10.1121/1.1907764.
- E. A. D. Clifton, J. R. B. Perry, F. Imamura, L. A. Lotta, S. Brage, N. G. Forouhi, S. J. Griffin, N. J. Wareham, K. K. Ong, and F. R. Day. Genome-wide association study for risk taking propensity indicates shared pathways with body mass index. *Communications Biology*, 1(1):36, Dec. 2018. ISSN 2399-3642. doi: 10.1038/s42003-018-0042-6.

- T. Cogley and T. J. Sargent. Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making\*. *International Economic Review*, 49(1):185–221, 2008. ISSN 1468-2354. doi: 10.1111/j.1468-2354.2008.00477.x.
- L. D. Cohn, S. Macfarlane, C. Yanez, and W. K. Imai. Risk-perception: Differences between adolescents and adults. *Health Psychology*, 14(3):217–222, 1995. ISSN 1930-7810, 0278-6133. doi: 10.1037/0278-6133.14.3.217.
- F. Condé, E. Maire-lepoivre, E. Audinat, and F. Crépel. Afferent connections of the medial frontal cortex of the rat. II. Cortical and subcortical afferents. *Journal of Comparative Neurology*, 352(4):567–593, 1995. ISSN 1096-9861. doi: 10.1002/cne.903520407.
- C. M. Constantinople, A. T. Piet, P. Bibawi, A. Akrami, C. Kopec, and C. D. Brody. Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *eLife*, 8:e49744, Nov. 2019a. ISSN 2050-084X. doi: 10.7554/eLife.49744.
- C. M. Constantinople, A. T. Piet, and C. D. Brody. An Analysis of Decision under Risk in Rats. *Current Biology*, 29(12):2066–2074.e5, June 2019b. ISSN 0960-9822. doi: 10.1016/j.cub.2019.05.013.
- S. E. Cooper, S. P. Goings, J. Y. Kim, and R. I. Wood. Testosterone enhances risk tolerance without altering motor impulsivity in male rats. *Psychoneuroendocrinology*, 40:201–212, Feb. 2014. ISSN 0306-4530. doi: 10.1016/j.psyneuen.2013.11.017.
- J. V. Corwin and R. L. Reep. Rodent posterior parietal cortex as a component of a cortical network mediating directed spatial attention. *Psychobiology*, 26(2):87–102, June 1998. ISSN 0889-6313. doi: 10.3758/BF03330596.

P. Dayan and A. 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. 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, Oct. 2018.

K. Deisseroth. Circuit dynamics of adaptive and maladaptive behaviour. *Nature*, 505(7483):309–317, Jan. 2014. ISSN 1476-4687. doi: 10.1038/nature12982.

M. L. Dent, L. A. Screven, and A. Kobra. Hearing in Rodents. In M. L. Dent, R. R. Fay, and A. 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.

A. Diedrich. A further test of sequential-sampling models that account for payoff effects on response bias in perceptual decision tasks. *Perception & Psychophysics*, 70(2):229–256, Feb. 2008. ISSN 1532-5962. doi: 10.3758/PP.70.2.229.

A. Diedrich and J. R. Busemeyer. Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, 68(2):194–207, Feb. 2006. ISSN 1532-5962. doi: 10.3758/BF03193669.

M. C. Dorris and P. W. Glimcher. Activity in Posterior Parietal Cortex Is Correlated with the Relative Subjective Desirability of Action. *Neuron*, 44(2):365–378, Oct. 2004. ISSN 08966273. doi: 10.1016/j.neuron.2004.09.009.

- C. A. Duan, J. C. Erlich, and C. D. Brody. Requirement of Prefrontal and Midbrain Regions for Rapid Executive Control of Behavior in the Rat. *Neuron*, 86(6):1491–1503, June 2015. ISSN 08966273. doi: 10.1016/j.neuron.2015.05.042.
- C. L. Ebbesen, M. N. Insanally, C. D. Kopec, M. Murakami, A. Saiki, and J. C. Erlich. More than Just a “Motor”: Recent Surprises from the Frontal Cortex. *Journal of Neuroscience*, 38(44):9402–9413, Oct. 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1671-18.2018.
- C. C. Eckel and P. 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 B. Y. Hayden. Risk aversion in macaques in a freely moving patch-leaving foraging task. Preprint, Animal Behavior and Cognition, Oct. 2018.
- D. Ellsberg. Risk, Ambiguity, and the Savage Axioms\*. *The Quarterly Journal of Economics*, 75(4):643–669, Nov. 1961. ISSN 0033-5533. doi: 10.2307/1884324.
- I. Erev and G. Barron. On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, 112(4):912–931, Oct. 2005. ISSN 0033-295X. doi: 10.1037/0033-295X.112.4.912.
- J. C. Erlich, M. Bialek, and C. D. Brody. A Cortical Substrate for Memory-Guided Orienting in the Rat. *Neuron*, 72(2):330–343, Oct. 2011. ISSN 0896-6273. doi: 10.1016/j.neuron.2011.07.010.
- J. C. Erlich, B. W. Brunton, C. A. Duan, T. D. Hanks, and C. 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.

J. Essig, J. B. Hunt, and G. Felsen. Inhibitory neurons in the superior colliculus mediate selection of spatially-directed movements. *Communications Biology*, 4(1): 719, Dec. 2021. ISSN 2399-3642. doi: 10.1038/s42003-021-02248-1.

S. Farashahi, H. Azab, B. Hayden, and A. 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.

S. Farashahi, C. H. Donahue, B. Y. Hayden, D. Lee, and A. Soltani. Flexible combination of reward information across primates. *Nature Human Behaviour*, 3(11): 1215–1224, Nov. 2019. ISSN 2397-3374. doi: 10.1038/s41562-019-0714-3.

S. Feng, P. Holmes, A. Rorie, and W. T. Newsome. Can Monkeys Choose Optimally When Faced with Noisy Stimuli and Unequal Rewards? *PLOS Computational Biology*, 5(2):e1000284, Feb. 2009. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1000284.

T. H. B. FitzGerald, B. Seymour, and R. J. Dolan. The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects. *Journal of Neuroscience*, 29(26):8388–8395, July 2009. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0717-09.2009.

S. B. Floresco, D. R. Montes, M. M. Tse, and M. van Holstein. Differential Contributions of Nucleus Accumbens Subregions to Cue-Guided Risk/Reward Decision Making and Implementation of Conditional Rules. *The Journal of Neuroscience*, 38

(8):1901–1914, Feb. 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3191-17.2018.

A. Frazzini. The Disposition Effect and Underreaction to News. *The Journal of Finance*, 61(4):2017–2046, 2006. ISSN 1540-6261. doi: 10.1111/j.1540-6261.2006.00896.x.

S. J. Galvin, J. V. Podd, V. Drga, and J. Whitmore. Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, 10(4):843–876, Dec. 2003. ISSN 1531-5320. doi: 10.3758/BF03196546.

J. Gao, R. Tortell, and J. L. McClelland. Dynamic Integration of Reward and Stimulus Information in Perceptual Decision-Making. *PLOS ONE*, 6(3):e16749, Mar. 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0016749.

M. P. Gardner, J. C. Conroy, C. V. Styer, T. Huynh, L. R. Whitaker, and G. Schoenbaum. Medial orbitofrontal inactivation does not affect economic choice. *eLife*, 7:e38963, Oct. 2018. ISSN 2050-084X. doi: 10.7554/eLife.38963.

M. P. H. Gardner, J. S. Conroy, M. H. Shaham, C. V. Styer, and G. Schoenbaum. Lateral Orbitofrontal Inactivation Dissociates Devaluation-Sensitive Behavior and Economic Choice. *Neuron*, 96(5):1192–1203.e4, Dec. 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2017.10.026.

M. P. H. Gardner, J. C. Conroy, D. C. Sanchez, J. Zhou, and G. Schoenbaum. Real-Time Value Integration during Economic Choice Is Regulated by Orbitofrontal

Cortex. *Current Biology*, 29(24):4315–4322.e4, Dec. 2019. ISSN 0960-9822. doi: 10.1016/j.cub.2019.10.058.

M. J. Goard, G. N. Pho, J. Woodson, and M. Sur. Distinct roles of visual, parietal, and frontal motor cortices in memory-guided sensorimotor decisions. *eLife*, 5:e13764, Aug. 2016. ISSN 2050-084X. doi: 10.7554/eLife.13764.

J. I. Gold, C.-T. Law, P. Connolly, and S. Bennur. The Relative Influences of Priors and Sensory Evidence on an Oculomotor Decision Variable During Perceptual Learning. *Journal of Neurophysiology*, 100(5):2653–2668, Nov. 2008. ISSN 0022-3077. doi: 10.1152/jn.90629.2008.

B. Hangya, J. I. Sanders, and A. Kepecs. A Mathematical Framework for Statistical Decision Confidence. *Neural Computation*, 28(9):1840–1858, Sept. 2016. ISSN 0899-7667. doi: 10.1162/NECO\_a\_00864.

T. D. Hanks and C. Summerfield. Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93(1):15–31, Jan. 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.12.003.

T. D. Hanks, C. D. Kopec, B. W. Brunton, C. A. Duan, J. C. Erlich, and C. D. Brody. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520(7546):220–223, Apr. 2015. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature14066.

T. A. Hare, J. O'Doherty, C. F. Camerer, W. Schultz, and A. 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.

K. M. Harlé, J. J. B. Allen, and A. G. Sanfey. The impact of depression on social economic decision making. *Journal of Abnormal Psychology*, 119(2):440–446, 2010. ISSN 1939-1846(Electronic),0021-843X(Print). doi: 10.1037/a0018612.

L. M. Harrison, A. Duggins, and K. J. Friston. Encoding uncertainty in the hippocampus. *Neural Networks*, 19(5):535–546, June 2006. ISSN 0893-6080. doi: 10.1016/j.neunet.2005.11.002.

J. Hartog, A. Ferrer-i-Carbonell, and N. Jonker. Linking Measured Risk Aversion to Individual Characteristics. *Kyklos*, 55(1):3–26, 2002. ISSN 1467-6435. doi: 10.1111/1467-6435.00175.

C. D. Harvey, P. Coen, and D. W. Tank. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*, 484(7392):62–68, Apr. 2012. ISSN 1476-4687. doi: 10.1038/nature10918.

J. Haushofer and E. Fehr. On the psychology of poverty. *Science*, 344(6186):862–867, May 2014. doi: 10.1126/science.1232491.

B. Hayden and M. Platt. Neurons in anterior cingulate cortex multiplex information about reward and action. *Journal of Neuroscience*, 30(9):3339–3346, 2010. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.4874-09.2010.

B. Y. Hayden and M. L. Platt. Temporal Discounting Predicts Risk Sensitivity in Rhesus Macaques. *Current Biology*, 17(1):49–53, Jan. 2007. ISSN 0960-9822. doi: 10.1016/j.cub.2006.10.055.

- C. Heath and A. Tversky. Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1):5–28, Jan. 1991. ISSN 1573-0476. doi: 10.1007/BF00057884.
- E. A. Heerey, K. R. Bell-Warren, and J. M. Gold. Decision-making impairments in the context of intact reward sensitivity in schizophrenia. *Biological Psychiatry*, 64(1):62–69, July 2008. ISSN 1873-2402. doi: 10.1016/j.biopsych.2008.02.015.
- S. Heilbronner, B. Hayden, and M. Platt. Decision Salience Signals in Posterior Cingulate Cortex. *Frontiers in Neuroscience*, 5:55, 2011. ISSN 1662-453X. doi: 10.3389/fnins.2011.00055.
- S. R. Heilbronner. Modeling risky decision-making in nonhuman animals: Shared core features. *Current Opinion in Behavioral Sciences*, 16:23–29, Aug. 2017. ISSN 2352-1546. doi: 10.1016/j.cobeha.2017.03.001.
- S. R. Heilbronner and B. 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.
- S. R. Heilbronner and B. Y. Hayden. The description-experience gap in risky choice in nonhuman primates. *Psychonomic Bulletin & Review*, 23(2):593–600, Apr. 2016. ISSN 1531-5320. doi: 10.3758/s13423-015-0924-2.
- R. Hertwig and I. Erev. The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12):517–523, Dec. 2009. ISSN 1364-6613. doi: 10.1016/j.tics.2009.09.004.

- R. Hertwig, G. Barron, E. U. Weber, and I. Erev. Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, 15(8):534–539, Aug. 2004. ISSN 0956-7976. doi: 10.1111/j.0956-7976.2004.00715.x.
- R. Hertwig, R. M. Hogarth, and T. Lejarraga. Experience and Description: Exploring Two Paths to Knowledge. *Current Directions in Psychological Science*, 27(2):123–128, Apr. 2018. ISSN 0963-7214. doi: 10.1177/0963721417740645.
- T. Hill, P. Kusev, and P. van Schaik. Choice Under Risk: How Occupation Influences Preferences. *Frontiers in Psychology*, 10:2003, 2019. ISSN 1664-1078. doi: 10.3389/fpsyg.2019.02003.
- D. Hocker, C. D. Brody, C. Savin, and C. M. Constantinople. Subpopulations of neurons in lOFC encode previous and current rewards at time of choice, May 2021.
- W. F. Hoffman, D. L. Schwartz, M. S. Huckans, B. H. McFarland, G. Meiri, A. A. Stevens, and S. H. Mitchell. Cortical activation during delay discounting in abstinent methamphetamine dependent individuals. *Psychopharmacology*, 201(2):183–193, Dec. 2008. ISSN 0033-3158. doi: 10.1007/s00213-008-1261-1.
- C. A. Holt, S. K. Laury, et al. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.
- S. A. Huettel, C. J. Stowe, E. M. Gordon, B. T. Warner, and M. L. Platt. Neural signatures of economic preferences for risk and ambiguity. *Neuron*, 49(5):765–775, Mar. 2006. ISSN 0896-6273. doi: 10.1016/j.neuron.2006.01.024.
- L. E. Hunter and S. J. Gershman. Reference-dependent preferences arise from structure learning, Nov. 2018.

E. J. Hwang, J. E. Dahlen, M. Mukundan, and T. Komiyama. History-based action selection bias in posterior parietal cortex. *Nature Communications*, 8(1):1242, Nov. 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-01356-z.

M. N. Insanally, I. Carcea, R. E. Field, C. C. Rodgers, B. DePasquale, K. Rajan, M. R. DeWeese, B. F. Albanna, and R. C. Froemke. Spike-timing-dependent ensemble encoding by non-classically responsive cortical neurons. *eLife*, 8:e42409, Jan. 2019. ISSN 2050-084X. doi: 10.7554/eLife.42409.

N. L. Jenni, J. D. Larkin, and S. B. Floresco. Prefrontal Dopamine D1 and D2 Receptors Regulate Dissociable Aspects of Decision Making via Distinct Ventral Striatal and Amygdalar Circuits. *Journal of Neuroscience*, 37(26):6200–6213, June 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0030-17.2017.

J. W. Kable and P. W. Glimcher. The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10(12):1625–1633, Dec. 2007. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn2007.

A. Kacelnik and M. Bateson. Risky Theories—The Effects of Variance on Foraging Decisions. *Integrative and Comparative Biology*, 36(4):402–434, Sept. 1996. ISSN 1540-7063. doi: 10.1093/icb/36.4.402.

A. Kacelnik and F. Brito e Abreu. Risky Choice and Weber's Law. *Journal of Theoretical Biology*, 194(2):289–298, Sept. 1998. ISSN 0022-5193. doi: 10.1006/jtbi.1998.0763.

A. Kacelnik and C. El Mouden. Triumphs and trials of the risk paradigm. *Animal*

*Behaviour*, 86(6):1117–1129, Dec. 2013. ISSN 0003-3472. doi: 10.1016/j.anbehav. 2013.09.034.

- D. Kahneman and S. 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.
- D. Kahneman and A. Tversky. Prospect Theory: An Analysis of Decision Under Risk. In *Handbook of the Fundamentals of Financial Decision Making*, volume Volume 4 of *World Scientific Handbook in Financial Economics Series*, pages 99–127. WORLD SCIENTIFIC, June 1979. ISBN 978-981-4417-34-1. doi: 10.1142/9789814417358\_0006.
- D. Kahneman, J. L. Knetsch, and R. H. Thaler. Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, 98(6):1325–1348, Dec. 1990. ISSN 0022-3808. doi: 10.1086/261737.
- D. Kahneman, D. Kahneman, A. Tversky (eds, and D. Kahneman. 2Experienced Utility and Objective Happiness: A Moment-Based Approach, 1999.
- N. Kandasamy, B. Hardy, L. Page, M. Schaffner, J. Graggaber, A. S. Powlson, P. C. Fletcher, M. Gurnell, and J. Coates. Cortisol shifts financial risk preferences. *Proc Natl Acad Sci U S A*, 111(9):3608–3613, Mar. 2014. doi: 10.1073/pnas.1317908111.
- L. N. Katz, J. L. Yates, J. W. Pillow, and A. C. Huk. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, 535(7611):285–288, July 2016. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature18617.

- Y. Kawamura, T. Takahashi, X. Liu, N. Nishida, Y. Noda, A. Yoshikawa, T. Umekage, and T. Sasaki. Variation in the *DRD2* gene affects impulsivity in intertemporal choice. 2013, Jan. 2013. doi: 10.4236/ojpsych.2013.31005.
- S. W. Kennerley, A. F. Dahmubed, A. H. Lara, and J. D. Wallis. Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *Journal of Cognitive Neuroscience*, 21(6):1162–1178, June 2009. ISSN 0898-929X. doi: 10.1162/jocn.2009.21100.
- A. Kepcs and Z. F. Mainen. A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1594):1322–1337, May 2012. ISSN 1471-2970. doi: 10.1098/rstb.2012.0037.
- A. Kepcs, N. Uchida, H. A. Zariwala, and Z. F. Mainen. Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210):227–231, Sept. 2008. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature07200.
- A. Klos, E. U. Weber, and M. Weber. Investment Decisions and Time Horizon: Risk Perception and Risk Behavior in Repeated Gambles. *Management Science*, 51(12): 1777–1790, Dec. 2005. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.1050.0429.
- F. H. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin, 1921.
- K. Kobayashi and M. Hsu. Neural mechanisms of updating under reducible and irreducible uncertainty. *Journal of Neuroscience*, pages 0535–17, June 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0535-17.2017.

- S. Koot, A. Baars, P. Hesseling, R. van den Bos, and M. Joëls. Time-dependent effects of corticosterone on reward-based decision-making in a rodent model of the Iowa Gambling Task. *Neuropharmacology*, 70:306–315, July 2013. ISSN 0028-3908. doi: 10.1016/j.neuropharm.2013.02.008.
- C. D. Kopec, J. C. Erlich, B. W. Brunton, K. Deisseroth, and C. D. Brody. Cortical and Subcortical Contributions to Short- Term Memory for Orienting Movements. *Neuron*, 88(2):367–377, Oct. 2015. doi: 10.1016/j.neuron.2015.08.033.
- K. P. Körding and D. M. Wolpert. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10(7):319–326, July 2006. ISSN 13646613. doi: 10.1016/j.tics.2006.05.003.
- J. D. Kralik, E. R. Xu, E. J. Knight, S. A. Khan, and W. J. Levine. When Less Is More: Evolutionary Origins of the Affect Heuristic. *PLOS ONE*, 7(10):e46240, Oct. 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0046240.
- R. H. Kramer, A. Mourot, and H. Adesnik. Optogenetic pharmacology for control of native neuronal signaling proteins. *Nature Neuroscience*, 16(7):816–823, July 2013. ISSN 1546-1726. doi: 10.1038/nn.3424.
- A. Lak, G. M. Costa, E. Romberg, A. A. Koulakov, Z. F. Mainen, and A. Kepecs. Orbitofrontal Cortex Is Required for Optimal Waiting Based on Decision Confidence. *Neuron*, 84(1):190–201, Oct. 2014. ISSN 0896-6273. doi: 10.1016/j.neuron.2014.08.039.
- A. Lak, M. Okun, M. M. Moss, H. Gurnani, K. Farrell, M. J. Wells, C. B. Reddy, A. Kepecs, K. D. Harris, and M. Carandini. Dopaminergic and Prefrontal Basis of

Learning from Sensory Confidence and Reward Value. *Neuron*, 105(4):700–711.e6, Feb. 2020. ISSN 0896-6273. doi: 10.1016/j.neuron.2019.11.018.

J. D. Larkin, N. L. Jenni, and S. B. Floresco. Modulation of risk/reward decision making by dopaminergic transmission within the basolateral amygdala. *Psychopharmacology*, 233(1):121–136, Jan. 2016. ISSN 0033-3158, 1432-2072. doi: 10.1007/s00213-015-4094-8.

M. Lauriola and I. P. Levin. Relating individual differences in Attitude toward Ambiguity to risky choices. *Journal of Behavioral Decision Making*, 14(2):107–122, 2001. ISSN 1099-0771. doi: 10.1002/bdm.368.

D. Lee. Decision Making: From Neuroscience to Psychiatry. *Neuron*, 78(2):233–248, Apr. 2013. ISSN 08966273. doi: 10.1016/j.neuron.2013.04.008.

I. P. Levin, M. A. Snyder, and D. 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, Mar. 1988. ISSN 0022-3980. doi: 10.1080/00223980.1988.9712703.

D. J. Levy and P. W. Glimcher. The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, 22(6):1027–1038, Dec. 2012. ISSN 0959-4388. doi: 10.1016/j.conb.2012.06.001.

H. Levy and H. M. Markowitz. Approximating Expected Utility by a Function of Mean and Variance. *The American Economic Review*, 69(3):308–317, 1979. ISSN 0002-8282.

- I. Levy, J. Snell, A. J. Nelson, A. Rustichini, and P. W. Glimcher. Neural Representation of Subjective Value Under Risk and Ambiguity. *Journal of Neurophysiology*, 103(2):1036–1047, Feb. 2010. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00853.2009.
- A. M. Licata, M. T. Kaufman, D. Raposo, M. B. Ryan, J. P. Sheppard, and A. K. Churchland. Posterior Parietal Cortex Guides Visual Decisions in Rats. *The Journal of Neuroscience*, 37(19):4954–4966, May 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0105-17.2017.
- L. 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.
- E. A. Ludvig and M. 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.
- E. Lukinova, Y. Wang, S. F. Lehrer, and J. C. Erlich. Time preferences are reliable across time-horizons and verbal versus experiential tasks. *eLife*, 8:e39656, Feb. 2019. ISSN 2050-084X. doi: 10.7554/eLife.39656.
- W. J. Ma. Bayesian Decision Models: A Primer. *Neuron*, 104(1):164–175, Oct. 2019. ISSN 0896-6273. doi: 10.1016/j.neuron.2019.09.037.
- C. K. Machens. Flexible Control of Mutual Inhibition: A Neural Model of Two-Interval Discrimination. *Science*, 307(5712):1121–1124, Feb. 2005. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1104171.

- J. K. Maner, J. A. Richey, K. Cromer, M. Mallott, C. W. Lejuez, T. E. Joiner, and N. B. Schmidt. Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences*, 42(4):665–675, Mar. 2007. ISSN 01918869. doi: 10.1016/j.paid.2006.08.016.
- J. G. March. Learning to be risk averse. *Psychological Review*, 103(2):309–319, Apr. 1996. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.103.2.309.
- H. Markowitz. The Utility of Wealth. *Journal of Political Economy*, 60(2):151–158, Apr. 1952. ISSN 0022-3808, 1537-534X. doi: 10.1086/257177.
- H. Markowitz. Mean–variance approximations to expected utility. *European Journal of Operational Research*, 234(2):346–355, Apr. 2014. ISSN 0377-2217. doi: 10.1016/j.ejor.2012.08.023.
- H. M. Markowitz. *Portfolio Selection*. Yale University Press, Jan. 1968. ISBN 978-0-300-19167-7.
- B. Marsh. Do Animals Use Heuristics? *Journal of Bioeconomics*, 4(1):49–56, Jan. 2002. ISSN 1573-6989. doi: 10.1023/A:1020655022163.
- B. Marsh and A. Kacelnik. Framing effects and risky decisions in starlings. *Proceedings of the National Academy of Sciences*, 99(5):3352–3355, Mar. 2002. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.042491999.
- A. T. Marshall and K. 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.

- A. N. McCoy and M. L. Platt. Expectations and outcomes: Decision-making in the primate brain. *Journal of Comparative Physiology A*, 191(3):201–211, Mar. 2005. ISSN 0340-7594, 1432-1351. doi: 10.1007/s00359-004-0565-9.
- K. J. Miller, M. M. Botvinick, and C. D. Brody. Dorsal hippocampus contributes to model-based planning. *Nature Neuroscience*, 20(9):1269–1276, Sept. 2017. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4613.
- J. Moisander, A. Markkula, and K. Eräranta. Construction of consumer choice in the market: Challenges for environmental policy. *International Journal of Consumer Studies*, 34(1):73–79, 2010. ISSN 1470-6431. doi: 10.1111/j.1470-6431.2009.00821.x.
- I. E. Monosov, D. A. Leopold, and O. Hikosaka. Neurons in the Primate Medial Basal Forebrain Signal Combined Information about Reward Uncertainty, Value, and Punishment Anticipation. *Journal of Neuroscience*, 35(19):7443–7459, May 2015. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0051-15.2015.
- I. E. Monosov, S. N. Haber, E. C. Leuthardt, and A. Jezzini. Anterior Cingulate Cortex and the Control of Dynamic Behavior in Primates. *Current Biology*, 30(23):R1442–R1454, Dec. 2020. ISSN 0960-9822. doi: 10.1016/j.cub.2020.10.009.
- A. S. Morcos and C. D. Harvey. History-dependent variability in population dynamics during evidence accumulation in cortex. *Nature Neuroscience*, 19(12):1672–1681, Dec. 2016. ISSN 1546-1726. doi: 10.1038/nn.4403.
- M. J. Mulder, E.-J. Wagenmakers, R. Ratcliff, W. Boekel, and B. U. Forstmann. Bias in the brain: A diffusion model analysis of prior probability and potential payoff.

*The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(7):2335–2343, Feb. 2012. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.4156-11. 2012.

M. R. Nassar, R. C. Wilson, B. Heasly, and J. I. Gold. An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. *Journal of Neuroscience*, 30(37):12366–12378, Sept. 2010. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0822-10.2010.

H. Neth and G. Gigerenzer. Heuristics: Tools for an Uncertain World. page 18.

W. T. Newsome and E. B. Pare. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, 8(6): 2201–2211, June 1988. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.08-06-02201.1988.

N. Nikbakht, A. Tafreshiha, D. Zoccolan, and M. E. Diamond. Supralinear and supramodal integration of visual and tactile signals in rats: Psychophysics and neuronal mechanisms. *Neuron*, 97(3):626–639, 2018.

Y. Niv, D. Joel, I. Meilijson, and E. Ruppin. Evolution of Reinforcement Learning in Uncertain Environments: A Simple Explanation for Complex Foraging Behaviors. *Adaptive Behavior*, 10(1):5–24, Jan. 2002. ISSN 1059-7123, 1741-2633. doi: 10.1177/1059-712302-010001-01.

Y. Niv, J. A. Edlund, P. Dayan, and J. P. O'Doherty. Neural Prediction Errors Reveal a Risk-Sensitive Reinforcement-Learning Process in the Human Brain. *Journal of*

*Neuroscience*, 32(2):551–562, Jan. 2012. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.5498-10.2012.

T. O'Connor and C. Franklin. Free Will. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2021 edition, 2021.

J. P. O'Doherty, S. W. Lee, and D. McNamee. The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, 1:94–100, Feb. 2015. ISSN 23521546. doi: 10.1016/j.cobeha.2014.10.004.

M. Ogawa, M. A. A. van der Meer, G. R. Esber, D. H. Cerri, T. A. Stalnaker, and G. Schoenbaum. Risk-Responsive Orbitofrontal Neurons Track Acquired Salience. *Neuron*, 77(2):251–258, Jan. 2013. ISSN 0896-6273. doi: 10.1016/j.neuron.2012.11.006.

M. O'Neill and W. Schultz. Coding of Reward Risk by Orbitofrontal Neurons Is Mostly Distinct from Coding of Reward Value. *Neuron*, 68(4):789–800, Nov. 2010. ISSN 0896-6273. doi: 10.1016/j.neuron.2010.09.031.

J. R. S. Onge, H. Abhari, and S. B. Floresco. Dissociable Contributions by Prefrontal D1 and D2 Receptors to Risk-Based Decision Making. *Journal of Neuroscience*, 31(23):8625–8633, June 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1020-11.2011.

C. A. Orsini, D. E. Moorman, J. W. Young, B. Setlow, and S. B. Floresco. Neural mechanisms regulating different forms of risk-related decision-making: Insights

from animal models. *Neuroscience & Biobehavioral Reviews*, 58:147–167, Nov. 2015. ISSN 01497634. doi: 10.1016/j.neubiorev.2015.04.009.

C. A. Orsini, C. M. Hernandez, S. Singhal, K. B. Kelly, C. J. Frazier, J. L. Bizon, and B. Setlow. Optogenetic Inhibition Reveals Distinct Roles for Basolateral Amygdala Activity at Discrete Time Points during Risky Decision Making. *The Journal of Neuroscience*, 37(48):11537–11548, Nov. 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2344-17.2017.

T. Pachur, R. Hertwig, G. Gigerenzer, and E. 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.

C. Padoa-Schioppa. Neurobiology of economic choice: A good-based model. *Annual Review of Neuroscience*, 34:333–359, 2011. ISSN 1545-4126. doi: 10.1146/annurev-neuro-061010-113648.

C. Padoa-Schioppa and J. A. Assad. Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441(7090):223–226, May 2006. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature04676.

C. Padoa-Schioppa and J. A. Assad. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature neuroscience*, 11(1): 95–102, Jan. 2008. ISSN 1097-6256. doi: 10.1038/nn2020.

E. Payzan-LeNestour, S. Dunne, P. Bossaerts, and J. 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.

G. N. Pho, M. J. Goard, J. Woodson, B. Crawford, and M. Sur. Task-dependent representations of stimulus and choice in mouse parietal cortex. *Nature Communications*, 9(1):2596, July 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05012-y.

A. T. Piet, J. C. Erlich, C. D. Kopec, and C. D. Brody. Rat Prefrontal Cortex Inactivations during Decision Making Are Explained by Bistable Attractor Dynamics. *Neural Computation*, 29(11):2861–2886, Aug. 2017. ISSN 0899-7667. doi: 10.1162/neco\_a\_01005.

S. Pisupati, L. Chartarifsky-Lynn, A. Khanal, and A. K. Churchland. Lapses in perceptual judgments reflect exploration. Preprint, Neuroscience, Apr. 2019.

S. Pisupati, L. Chartarifsky-Lynn, A. Khanal, and A. K. Churchland. Lapses in perceptual decisions reflect exploration. *eLife*, 10:e55490, Jan. 2021. ISSN 2050-084X. doi: 10.7554/eLife.55490.

M. L. Platt and P. W. Glimcher. Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741):233–238, July 1999. ISSN 0028-0836, 1476-4687. doi: 10.1038/22268.

M. L. Platt and S. A. Huettel. Risky business: The neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, 11(4):398–403, Apr. 2008. ISSN 1546-1726. doi: 10.1038/nn2062.

A. Pouget, J. Drugowitsch, and A. Kepcs. Confidence and certainty: Distinct prob-

abilistic quantities for different goals. *Nature Neuroscience*, 19(3):366–374, Mar. 2016. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4240.

J. W. Pratt. 4 - RISK AVERSION IN THE SMALL AND IN THE LARGE11This research was supported by the National Science Foundation (grant NSF-G24035). Reproduction in whole or in part is permitted for any purpose of the United States Government. In P. Diamond and M. Rothschild, editors, *Uncertainty in Economics*, pages 59–79. Academic Press, Jan. 1964. ISBN 978-0-12-214850-7. doi: 10.1016/B978-0-12-214850-7.50010-3.

K. Preuschoff, S. Quartz, and P. Bossaerts. Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, 28(11):2745–2752, 2008. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.4286-07.2008.

M. Rabin and R. H. Thaler. Anomalies: Risk Aversion. *Journal of Economic Perspectives*, 15(1):219–232, Feb. 2001. ISSN 0895-3309. doi: 10.1257/jep.15.1.219.

L.-L. Rao, Y. Zhou, D. Zheng, L.-Q. Yang, and S. 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, Oct. 2018. ISSN 0956-7976, 1467-9280. doi: 10.1177/0956797618779961.

D. Raposo, M. T. Kaufman, and A. K. Churchland. A category-free neural population supports evolving demands during decision-making. *Nature neuroscience*, 17(12): 1784–1792, Dec. 2014. ISSN 1097-6256. doi: 10.1038/nn.3865.

R. Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978. ISSN 1939-1471(Electronic),0033-295X(Print). doi: 10.1037/0033-295X.85.2.59.

R. L. Reep and J. V. Corwin. Posterior parietal cortex as part of a neural network for directed attention in rats. *Neurobiology of Learning and Memory*, 91(2):104–113, Feb. 2009. ISSN 1074-7427. doi: 10.1016/j.nlm.2008.08.010.

R. L. Reep, J. V. Corwin, A. Hashimoto, and R. T. Watson. Efferent Connections of the Rostral Portion of Medial Agranular Cortex in Rats. *Brain Research Bulletin*, 19(2):203–221, Aug. 1987. ISSN 0361-9230. doi: 10.1016/0361-9230(87)90086-4.

M. R. Roesch, A. R. Taylor, and G. Schoenbaum. Encoding of Time-Discounted Rewards in Orbitofrontal Cortex Is Independent of Value Representation. *Neuron*, 51(4):509–520, Aug. 2006. ISSN 0896-6273. doi: 10.1016/j.neuron.2006.06.027.

J. D. Roitman and M. N. Shadlen. Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *Journal of Neuroscience*, 22(21):9475–9489, Nov. 2002. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.22-21-09475.2002.

A. E. Rorie, J. Gao, J. L. McClelland, and W. T. Newsome. Integration of Sensory and Reward Information during Perceptual Decision-Making in Lateral Intraparietal Cortex (LIP) of the Macaque Monkey. *PLOS ONE*, 5(2):e9308, Feb. 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0009308.

R. B. Rutledge, P. Smittenaar, P. Zeidman, H. R. Brown, R. A. Adams, U. Lindenberger, P. Dayan, and R. 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.

R. A. Saez, A. Saez, J. J. Paton, B. Lau, and C. D. Salzman. Distinct Roles for

the Amygdala and Orbitofrontal Cortex in Representing the Relative Amount of Expected Reward. *Neuron*, 95(1):70–77.e3, July 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2017.06.012.

K. Samejima, Y. Ueda, K. Doya, and M. Kimura. Representation of action-specific reward values in the striatum. *Science (New York, N.Y.)*, 310(5752):1337–1340, Nov. 2005. ISSN 1095-9203. doi: 10.1126/science.1115270.

P. Samuelson. Risk and uncertainty: A fallacy of large numbers. *Scientia*, 57(98), 1963.

J. I. Sanders, B. Hangya, and A. Kepecs. Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron*, 90(3):499–506, May 2016. ISSN 08966273. doi: 10.1016/j.neuron.2016.03.025.

W. Schultz, K. Preuschoff, C. Camerer, M. Hsu, C. D. Fiorillo, P. N. Tobler, and P. Bossaerts. Explicit neural signals reflecting reward uncertainty. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1511):3801–3811, Dec. 2008. doi: 10.1098/rstb.2008.0152.

B. B. Scott, C. M. Constantinople, J. C. Erlich, D. W. Tank, and C. D. Brody. Sources of noise during accumulation of evidence in unrestrained and voluntarily head-restrained rats. *eLife*, 4:e11308, Dec. 2015. ISSN 2050-084X. doi: 10.7554/eLife.11308.

B. B. Scott, C. M. Constantinople, A. Akrami, T. D. Hanks, C. D. Brody, and D. W. Tank. Fronto-parietal Cortical Circuits Encode Accumulated Evidence with

a Diversity of Timescales. *Neuron*, 95(2):385–398.e5, July 2017. ISSN 08966273. doi: 10.1016/j.neuron.2017.06.013.

J. T. Serences. Value-Based Modulations in Human Visual Cortex. *Neuron*, 60(6): 1169–1181, Dec. 2008. ISSN 0896-6273. doi: 10.1016/j.neuron.2008.10.051.

G. Sescousse, X. Caldú, B. Segura, and J.-C. 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.

M. N. Shadlen and W. T. Newsome. Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *Journal of Neurophysiology*, 86 (4):1916–1936, Oct. 2001. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.2001.86.4. 1916.

S. Shafir, T. A. Waite, and B. 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, Jan. 2002. ISSN 1432-0762. doi: 10.1007/s00265-001-0420-8.

W. 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.

K. G. Shimp, M. R. Mitchell, B. S. Beas, J. L. Bizon, and B. Setlow. Affective and cognitive mechanisms of risky decision making. *Neurobiology of Learning and Memory*, 117:60–70, Jan. 2015. ISSN 1074-7427. doi: 10.1016/j.nlm.2014.03.002.

- J. Shires, S. Joshi, and M. A. Basso. Shedding new light on the role of the basal ganglia-superior colliculus pathway in eye movements. *Current Opinion in Neurobiology*, 20(6):717–725, Dec. 2010. ISSN 09594388. doi: 10.1016/j.conb.2010.08.008.
- N. W. Simon, R. J. Gilbert, J. D. Mayse, J. L. Bizon, and B. Setlow. Balancing Risk and Reward: A Rat Model of Risky Decision Making. *Neuropsychopharmacology*, 34(10):2208–2217, Sept. 2009. ISSN 1740-634X. doi: 10.1038/npp.2009.48.
- M. J. Siniscalchi, V. Phoumthipphavong, F. Ali, M. Lozano, and A. C. Kwan. Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior. *Nature Neuroscience*, 19(9):1234–1242, Sept. 2016. ISSN 1546-1726. doi: 10.1038/nn.4342.
- N.-Y. So and V. Stuphorn. Supplementary Eye Field Encodes Option and Action Value for Saccades With Variable Reward. *Journal of Neurophysiology*, 104(5):2634–2653, Nov. 2010. ISSN 0022-3077. doi: 10.1152/jn.00430.2010.
- A. Soltani and A. Izquierdo. Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, 20(10):635–644, Oct. 2019. ISSN 1471-0048. doi: 10.1038/s41583-019-0180-y.
- M. Spitmaan, E. Chu, and A. Soltani. Salience-Driven Value Construction for Adaptive Choice under Risk. *Journal of Neuroscience*, 39(26):5195–5209, June 2019. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2522-18.2019.
- J. R. St Onge and S. B. Floresco. Dopaminergic Modulation of Risk-Based Decision Making. *Neuropsychopharmacology*, 34(3):681–697, Feb. 2009. ISSN 1740-634X. doi: 10.1038/npp.2008.121.

L. Stănișor, C. van der Togt, C. M. A. Pennartz, and P. R. Roelfsema. A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences*, 110(22):9136–9141, May 2013. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1300117110.

W. R. Stauffer, A. Lak, and W. Schultz. Dopamine Reward Prediction Error Responses Reflect Marginal Utility. *Current Biology*, 24(21):2491–2500, Nov. 2014. ISSN 0960-9822. doi: 10.1016/j.cub.2014.08.064.

C. M. Stopper and S. 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, Mar. 2011. ISSN 1531-135X. doi: 10.3758/s13415-010-0015-9.

C. M. Stopper, E. B. Green, and S. B. Floresco. Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. *Cerebral cortex*, 24(1):154–162, 2012.

C. M. Stopper, S. Khayambashi, and S. B. Floresco. Receptor-Specific Modulation of Risk-Based Decision Making by Nucleus Accumbens Dopamine. *Neuropsychopharmacology*, 38(5):715–728, Apr. 2013. ISSN 1740-634X. doi: 10.1038/npp.2012.240.

J. A. Sugam, M. P. Saddoris, and R. 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.

J. H. Sul, S. Jo, D. Lee, and M. W. Jung. Role of rodent secondary motor cortex

in value-based action selection. *Nature Neuroscience*, 14(9):1202–1208, Sept. 2011. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.2881.

C. Summerfield and E. Koechlin. Economic Value Biases Uncertain Perceptual Choices in the Parietal and Prefrontal Cortices. *Frontiers in Human Neuroscience*, 4, 2010. ISSN 1662-5161. doi: 10.3389/fnhum.2010.00208.

D. G. Tervo, M. Proskurin, M. Manakov, M. Kabra, A. Vollmer, K. Branson, and A. Y. Karpova. Behavioral Variability through Stochastic Choice and Its Gating by Anterior Cingulate Cortex. *Cell*, 159(1):21–32, Sept. 2014. ISSN 00928674. doi: 10.1016/j.cell.2014.08.037.

P. N. Tobler and E. U. Weber. Valuation for Risky and Uncertain Choices. *Neuroeconomics: Decision Making and the Brain: Second Edition*, pages 149–172, Sept. 2013. doi: 10.1016/B978-0-12-416008-8.00009-7.

P. N. Tobler, G. I. Christopoulos, J. P. O'Doherty, R. J. Dolan, and W. Schultz. Risk-dependent reward value signal in human prefrontal cortex. *Proceedings of the National Academy of Sciences*, 106(17):7185–7190, Apr. 2009. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0809599106.

A. Tymula, L. A. R. Belmaker, L. Ruderman, P. W. Glimcher, and I. 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.

A. E. Urai, A. Braun, and T. H. Donner. Pupil-linked arousal is driven by decision

uncertainty and alters serial choice bias. *Nature Communications*, 8(1):14637, Mar. 2017. ISSN 2041-1723. doi: 10.1038/ncomms14637.

E. Van Duuren, G. Van Der Plasse, J. Lankelma, R. Joosten, M. Feenstra, and C. Pennartz. Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. *Journal of Neuroscience*, 29(28):8965–8976, 2009. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.0005-09.2009.

M. van Holstein and S. B. Floresco. Dissociable roles for the ventral and dorsal medial prefrontal cortex in cue-guided risk/reward decision making. *Neuropharmacology*, 45(4):683–693, Mar. 2020. ISSN 1740-634X. doi: 10.1038/s41386-019-0557-7.

S. M. C. van Osch, W. B. van den Hout, and A. M. Stiggebout. Exploring the Reference Point in Prospect Theory: Gambles for Length of Life. *Medical Decision Making*, 26(4):338–346, July 2006. ISSN 0272-989X. doi: 10.1177/0272989X06290484.

A. Vehtari, A. Gelman, and J. Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, Sept. 2017. ISSN 1573-1375. doi: 10.1007/s11222-016-9696-4.

V. Venkatraman, J. W. Payne, and S. 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, Nov. 2014. ISSN 0749-5978. doi: 10.1016/j.obhdp.2014.06.003.

J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*.

Princeton University press, Princeton (N.J.), 3rd ed edition, 1953. ISBN 978-0-691-04183-4.

B. J. Weber and G. 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.

E. U. Weber, S. Shafir, and A.-R. Blais. Predicting Risk Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation. *Psychological Review*, 111(2):430–445, 2004. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.111.2.430.

J. R. Whitlock. Posterior parietal cortex. *Current Biology*, 27(14):R691–R695, July 2017. ISSN 0960-9822. doi: 10.1016/j.cub.2017.06.007.

A. F. Williams. Teenage drivers: Patterns of risk. *Journal of Safety Research*, 34(1): 5–15, Jan. 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.

Y.-H. Xuan, S. Li, R. Tao, J. Chen, L.-L. Rao, X. T. Wang, and R. Zheng. Genetic and Environmental Influences on Gambling: A Meta-Analysis of Twin Studies. *Frontiers in Psychology*, 8:2121, Dec. 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.02121.

- J. Yacubian, T. Sommer, K. Schroeder, J. Gläscher, D. F. Braus, and C. Büchel. Subregions of the ventral striatum show preferential coding of reward magnitude and probability. *NeuroImage*, 38(3):557–563, Nov. 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2007.08.007.
- J. F. 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.
- E. Yechiam and J. R. Busemeyer. The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making*, 19(1):1–16, 2006. ISSN 1099-0771. doi: 10.1002/bdm.509.
- K. A. Zalocusky, C. Ramakrishnan, T. N. Lerner, T. J. Davidson, B. Knutson, and K. Deisseroth. Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature*, 531(7596):642–646, Mar. 2016. ISSN 0028-0836. doi: 10.1038/nature17400.
- F. D. Zeeb, P. J. J. Baarendse, L. J. M. J. Vanderschuren, and C. A. Winstanley. Inactivation of the prelimbic or infralimbic cortex impairs decision-making in the rat gambling task. *Psychopharmacology*, 232(24):4481–4491, Dec. 2015. ISSN 1432-2072. doi: 10.1007/s00213-015-4075-y.
- L. Zhong, Y. Zhang, C. A. Duan, J. Deng, J. Pan, and N.-l. Xu. Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization. *Nature Neuroscience*, 22(6):963–973, June 2019. ISSN 1546-1726. doi: 10.1038/s41593-019-0383-6.

X. Zhu, J. Moller-Mara, S. Dubroqua, C. Bao, and J. C. Erlich. Frontal but not parietal cortex is required for decisions under risk, Nov. 2021.