ADAPTIVE ECONOMICS:
A neuroethological approach to the study of preferences, biases, and choice

Philipe Mathurin Bujold
Hughes Hall, Cambridge
January 2020
This dissertation is submitted for the degree of Doctor of Philosophy

Preface

Declaration

This dissertation was completed within the Department of Physiology, Development and Neuroscience under the supervision of Professor Wolfram Schultz.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in this preface and specified in the text. In Chapter 2, the data from monkey C was collected in collaboration with Dr Simone Ferrari-Toniolo. In Chapter 4, sections of work were carried out in collaboration with Dr Simone Ferrari-Toniolo (specifically, the Machina Triangle analysis and the validation of our model). It should also be noted that the chapter was published in The Journal of Neuroscience as: Ferrari-Toniolo, S.*, Bujold, P. M.*, & Schultz, W. (2019). Probability distortion depends on choice sequence in rhesus monkeys.

This dissertation is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

Finally, this dissertation does not exceed the prescribed word limit for the Biology Degree Committee.

ADAPTIVE ECONOMICS:
A neuroethological approach to the study of preferences, biases, and choice

by Philipe Mathurin Bujold

Dissertation summary

A neuron's curse is that at every given time, with the information available to it, it must choose to either send a signal to its neighbouring cells or remain silent. It has evolved to be the optimal decision unit and, together with around 86 billion of its neighbours, the neuron keeps us alive, helps us cooperate, and allows us to successfully compete with others when resources get scarce. Yet, we, being collections of these neurons, still struggle to describe how these individual decision-makers support the broader process that is human decision-making.

Traditionally, decision theory has sought to understand human choices by relying more on mathematics than biology. This has led to the general assumption that decision-makers behave ‘as-if’ guided by mathematical rules and algorithms that are mostly static over time. In reality, however, decision-making relies on a brain that, due to its limited capacity, has evolved the ability for flexible and dynamic cognition.

The experiments presented in this thesis, build on dichotomies in human behaviour that cannot be explained by traditional economic models - first replicating these findings in rhesus macaques, then addressing the neurobiological algorithms that could reconcile these dichotomies. Specifically, I looked at the effects of different reward ranges, different levels of risk, and different experimental paradigms in shaping the way monkeys made choices. I demonstrate that, far from having the stable and fixed preferences prescribed by economic models, rhesus macaques appear to flexibly adapt their choice preferences in a way that optimizes their decision-making given their experience with the task at hand. I then elaborate on the neurobiological basis for preference adaptation, and show how incorporating simple, dynamic algorithms into economic choice models improves their predictive power.

Taken together, my results demonstrate the need for, and advantage of, integrating neuroethological thought into the current framework of decision theory.

Acknowledgements

First and foremost, I would like to thank Professor Wolfram Schultz for his ongoing support and supervision throughout my time in Cambridge. I would also like to thank all of the other members of the Schultz lab for their generous help and invaluable feedback over the years. In particular, I am extremely grateful to Dr Simone Ferrari-Toniolo with whom I collaborated closely and who helped me to 'hit the ground running' at the start of my PhD.

To Christina Thomson and Aled David, thank you for dedicating what seems like every moment of your lives to making sure our monkeys are kept healthy and happy. And to my PhD advisor Simon Laughlin, thank you for encouraging me to always go further - to look at every problem from multiple perspectives.

I also want to thank Dr Jan Zimmerman and Dr Kelly Diederen, for their guidance and unwavering support. And to Paul Glimcher and Kenway Louie, thank you for introducing me to the world of neuroeconomics all those years ago.

To Eva, Mark, Harum, Alexia, and Charlotte, thank you for your valuable feedback and comments during the process of writing of this dissertation. And to Sarah, Konstantin, Robert, Putu, Daniel, and Alaa, thank you for making life in the lab a lot more exciting.

Finally, none of this work would have been achievable without the love and support of my friends and family who encouraged me all along the way. Thank you for believing in me, and for putting up with the ups and downs of my doctoral years. To my friend and flatmate Sabrina, thank you for grounding me, and for making our little house in Cambridge feel like home. To my partner Joseph, I could not have asked for a better teammate. Thank you for supporting me when I needed strength, listening when I needed an ear, and grounding me when I ventured too deep into my own head. But most importantly, thank you for loving me, always. And to my mother, father, and sister; none of this would have been possible without your support. The journey from Sherbrooke to Cambridge was a long one - you were there for me every step of the way. Je vous aime.

Finally, while they spent their lives making decisions, they never truly had a choice:

My monkeys, Tigger, Trident, and Ugo, Thank you.

Contents

Chapter 1

A brief history of decision-making

How do we make decisions? Or for that matter, how does any animal make any decision? This fundamental question, and the focus of 'decision theory', has long been dissected by experts in the fields of economics, neuroscience, psychology, ecology and philosophy - each of these looking at decision-making from very different perspectives.

Universally (and throughout this thesis), research on decision-making almost always involves the use of models: imperfect[1] algorithms that, given sets of inputs (things like the costs, the benefits, or the uncertainty associated with a given choice), output quantifiable behavioural predictions. We can think of these as strategies, or sets of rules, that we can use to predict when one might choose apples over oranges, to describe the activity of individual neurons, or to quantify differences in the foraging behaviour of animals (in the same ways that we use models to predict and understand the weather). Models are decision theory's bread and butter, but the purpose of models also divides decision theory into different theoretical schools: the prescriptive, descriptive, and the - often ignored - predictive.

Normative (or prescriptive) decision theory, conventionally the realm of neoclassical economics, concerns itself with the development of choice models that capture behavioural ideals; that is, they output what a purely rational decision-maker ought to do (Stevens, 2010; Thaler, 2016). Normative models assume we behave 'as-if' following specific sets of rules and axioms when making choices, and they see any deviations between our behaviour and their predictions as irrational (a failure in optimizing our decisions). These theories *prescribe* rather than *describe* actual choices, and are less concerned with realism than they are with objective optimization. The reverse, where the emphasis is on realism and capturing true behaviour, is the branch of decision theory we call descriptive (see Schoemaker, 1982 for review; but see also Chase, Hertwig, & Gigerenzer, 1998; and McFall, 2015). This approach is the prime focus of behavioural economics - the merger of psychology, sociology, and economics – a subfield of economics that appeared in the latter half of the 20th century as a direct response to the descriptive failures of the normative approach. As we will cover in this chapter, descriptive models attempt to best 'fit' real behaviour in a way the normative approach does not; where prescriptive models ask 'what' are the choices that we make, the descriptive approach – with the help of psychology – goes further and

---

[1] The power of modelling lies in the fact that models always serve as simplifications of the real world. Models are thus, in that sense, imperfect. They allow us to ignore those variables that we consider uninformative, focusing solely on the variables that we believe are most relevant to the question we are trying to answer. This is put perhaps most poetically by the statistician George Box: "all models are wrong but some are useful" (1979).

asks 'why'. The third and final school, the one we call predictive, interweaves neuroscience, comparative psychology, and ecology with traditional economic thought – doing so by asking 'how' (and when) we might make the decisions prescribed or described by the other two schools (Fig. 1-1).

The interaction of these three approaches is decision theory's greatest strength, but the way in which they interact also highlights one of its biggest limitation (particularly as we search for decision-making 'in the brain'): both the descriptive and predictive approaches currently rely on normative economics. Economists observe people's choices, they then use their measures to develop mathematical algorithms that describe human behaviour; psychologists refine these models using cognitive principles that 'make sense' of the choices measured by economists; and neuroscientists, working to understand the brain, use these models to make predictions about the neural circuits that govern decision-making. Put simply, we use what 'ought to be' as a basis for what 'is' - in this case, 'as-if' models that wholly ignore the biological constraints of our physiological brain (Sugden, 1993; Volz & Gigerenzer, 2014). Every level of decision theory is shaped, in some way or another, by the unrealistic prescriptive models of neoclassical economics. Could we not, instead, incorporate biology from the start? Should we not, if the goal is to predict and understand choices, develop models we know the brain can support?

This thesis' **First Chapter**, through an overview of decision theory's timeline, milestones, and disagreements, outlines the challenge that the pervasiveness of normative ideals poses for the fields that then rely on predictive accuracy (namely, neuroscience, comparative psychology, and ecology). Next, **Chapters 2, 3, and 4** tackle the above in hopes of consolidating the decision-making fields: first by highlighting the predictive limitations of even today's most popular decision models, then by demonstrating how simple edits rooted in neurobiological thought can improve the predictive validity of these algorithms. The experiments presented in these chapters focus on changes in behaviours that are not accounted for by traditional choice models, doing so with a twist. If the goal is to improve the biological validity of choice models, a crucial step is to ensure they can also account for the behaviour of animals that share in our neurobiology – other primates, in particular. The work presented in this thesis therefore focuses on the economic behaviour (choice preferences, patterns, and biases) of our close cousin, the *Rhesus Macaque (*Macaca mulatta), and on one of our most defining primate characteristic: *adaptability* (Jones, 2005). The last two chapters, the **Discussion** and **Postface**, stand in direct answer to the problem posed in this first one (the lack of biological validity in decision theory) – they synthesize my experimental results, reconcile past incongruities in the field, and cover the latest trends in decision theory.

Though each of my thesis' subsequent chapters will feature their own, topic-specific reviews, I invite you to view the rest of this chapter as the premise of what is to come, and as the inspiration behind my research focus. In the words of the evolutionary biologists Theodosius Dobzhansky:

> *'Nothing in biology makes sense except in the light of evolution'* …

… and decision-making is inherently, undoubtedly, biological.

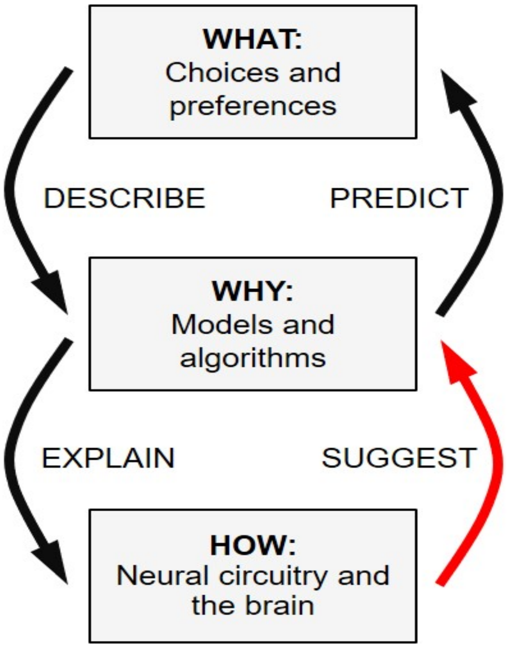


Figure 1-1 | The questions asked by descriptive decision-theory. The disciplines that study decision-making mostly do so by linking three different questions: *what*, *why*, and *how* do we make the choices we make. *What* questions focus on choices and preferences that can be observed in experimental settings. *Why* questions seek to explain the 'algorithms' that can lead to the patterns of choice we observe. *How* questions tie in these algorithms with biology – seeking to understand the neurobiological hardware that supports the choices we make. Though decision-theory has had much success moving downwards (from what to how questions), a crucial limitation to the modern framework of decision-theory is that no one tries to answer the questions in a bottom-up manner (from how to what). There is little research done where biology is the actual starting point. (For reviews on these questions and broader framework, see Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017; Marr & Ullman, 1982; Mobbs, Trimmer, Blumstein, & Dayan, 2018)

1.1 Prescription

For any and all disciplines studying decision-making, *rational choice theory* presents itself as the *de facto* starting point (von Neumann et al., 2004). The embodiment of the normative (prescriptive) approach, rational choice theory essentially states that decision-makers should act in a way that maximises their personal gains, doing so by ranking real-world outcomes onto a common scale (usually utility) to then select the choice that ranks highest. A rational decision-maker always chooses the option that they absolutely, and consistently prefer – the perfect optimizer. The present section illustrates how the framework of rational choice has come to dominate decision-making research, and how it led to the development of perhaps the single most important concept in decision theory: *utility maximisation*.

1.1.1 Expected Value

The idea behind rational choice theory can be traced back to the mid-seventeenth century, when a series of correspondences between the philosopher Blaise Pascal and the mathematician Pierre de Fermat led to the development of what is now known as modern probability theory (Ore, 1960). In these letters, the duo formalized the concept of *expected value* (EV), whereby the objective value of a choice could be expressed as the product of (i) its outcome value (or values) and (ii) the likelihood that this outcome would occur. Using this early definition of economic value, Pascal theorized that any rational decision-maker would make choices that maximize expected value returns. In its full form, the EV of a choice was (and still is) given by the sum of all its outcomes (x) multiplied by the respective probability (p) that they would come to pass:

$$\text{Expected Value} = \sum p_n \times x_n \tag{1-1}$$

Pascal and de Fermat's formalization of expected value linked economic theory with probability theory in two crucial ways: choices with the same mean outcome were recognized as having the same expected value (the EV corresponds to the statistical average, or mean, of a choice's outcome distribution); and the *variance* in choices' outcome distributions came to be known as the *risk* involved in selecting a given option. Thereafter, the term risk came to be used in decision theory to define the certainty with which a decision-maker can expect the outcome of a choice to reflect said choice's expected value.[2]

While EVs go a long way in predicting people's choices (see Fig. 1-2), a crucial limitation was made apparent in 1713 through a hypothetical situation that is now famously known as the St Petersburg paradox:

---

[2] This goes against the psychological, and indeed more popular use of risk as a proxy for a loss' magnitude. In economics, risk is always a positive number, regardless of choice outcomes being positive or negative, and relates specifically to the statistical concept of variance. It is also important to dissociate magnitude risk from probability risk; the variance in outcome magnitudes versus the variance in outcome likelihoods.

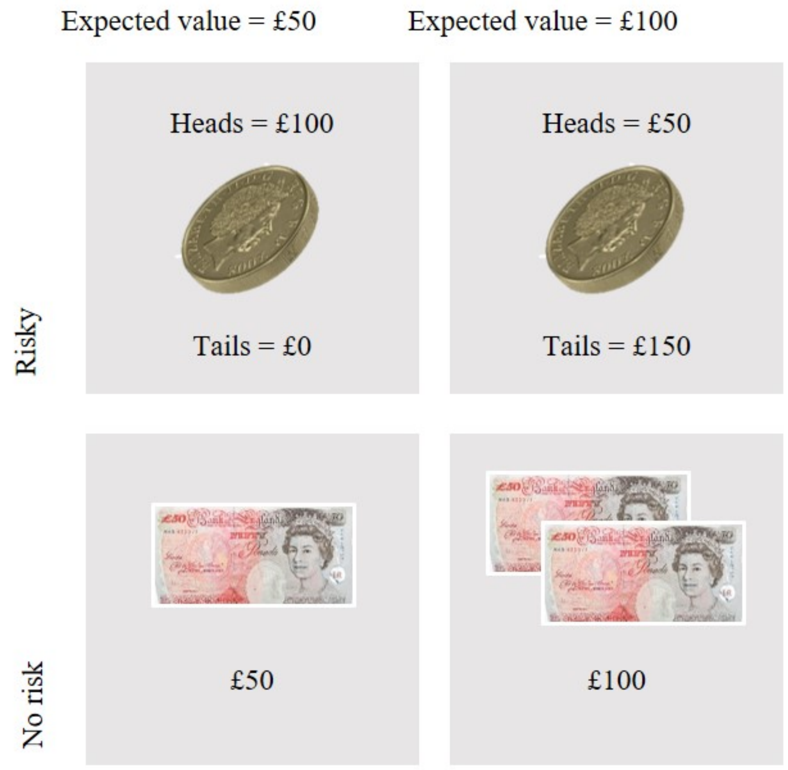


Figure 1-2 | Predicting choices from expected value. Four different choice options appear above. The first two, in the upper quadrant, represent coin flips: each outcome is uncertain, but their expected value is the sum of their individual outcomes' expected value. The lower two options represent what we call safe options: the outcome of the choice is guaranteed, and the expected value therefore the value of their single outcome. Each of these options is unique, but those that share a column are equal in expected value, whilst the options on the same line share an equal level of risk. Expected value predicts that the options on the right would be preferred to the options on the left.

> *"Suppose someone offers to toss a fair coin repeatedly until it comes up heads, and to pay you £1 if this happens on the first toss, £2 if it takes two tosses to land a head, £4 if it takes three tosses, £8 if it takes four tosses, and so on. What is the largest sure payment you would be willing to forgo in order to [...] play this game?"* (Machina, 1987, p.135)

If expected value is used to predict what a rational decision-maker would pay, we get a very puzzling prediction: because the game can – hypothetically - carry on forever, its expected value is infinite. That is, if rational decision-makers use EVs to make decisions, the rational price to pay for this game is the totality of one's wealth (as opposed to the few pounds any real person would truly pay):

$$\text{Expected Value } = \frac{1}{2} \times £1 + \frac{1}{4} \times £2 + \frac{1}{8} \times £4 + \cdots = £0.5 + £0.5 + £0.5 + \ldots = £\infty \quad (1\text{-}2)$$

This obvious discrepancy between real behaviour and the expected value predictions in the St Petersburg paradox led to the next development in rational choice theory: the *Expected Utility Hypothesis*.

1.1.2 Expected Utility Theory

In the eighteenth century, two centuries after Pascal and de Fermat's development of expected value, the Swiss Bernoulli family started exchanging letters on the topic of the St Petersburg paradox. Nicolas Bernoulli had first recognized the paradox in 1713 (Cowling, 1955), but it is his cousin, the mathematician Daniel Bernoulli, who finally presented a solution in 1738 (Bernoulli, 1954). Instead of making choices using expected values, Daniel Bernoulli proposed that people relied on their own, unique level of wealth in affording value to choices. This internal, subjective value came to be known as *utility*: the value that something has to you.

A £50 note is worth more to someone who has £0 than to someone who already has £100. Bernoulli reasoned that the metric used by decision-makers to make choices should reflect this change: the change in an outcome's 'worth' relative one's current level of wealth. According to his expected utility hypothesis, people would make choices in a way that maximized their subjective metric of utility (accounting for their context and tastes). People were utility-maximising rather than expected value maximising, and unlike expected value, the utility assigned to identical options could be different for different people. In Bernoulli's formulation, the expected utility of a choice stands as the sum of all individual outcome utilities ($u(x)$) multiplied by their respective probability (p):

$$\text{Expected Utility } = \sum p_n \times u(x_n) \quad (1\text{-}3)$$

The Bernoulli utility function, $u(x_n)$, represents the transformation of objective values (x_n) into utilities (a transformation that depends on the individual's preferences and their current level of wealth). Bernoulli theorised that this function takes the form of a gradually flattening upwards curve (a concave function). That is, as wealth increases, each additional unit of x_i results in less 'marginal' utility (Fig. 1-3).

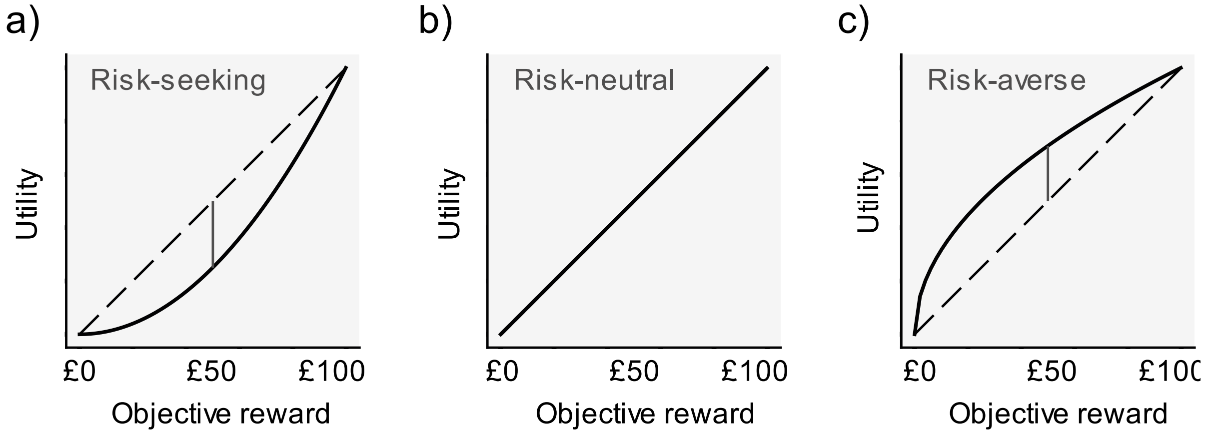


Figure 1-3 | Expected utility's curvature predicts risk attitude. Picture a choice between a coin flip that rewards £100 half the time (£0 otherwise), and a guaranteed £50 note – the choice one will make is a direct result of the utility function being a) convex, b) linear, or c) concave. In the case of convexity, the expected utility of the $50 note will always be lower than that of the coin flip (50% chance of £100, £0 otherwise). We say, in that case, that the decision-maker is risk-seeking and prefers riskier, more uncertain options. If we reverse the curvature and make the utility function concave, we find that the opposite is true: the coin flip will always have a lower utility than the safe $50 option. Finally, assuming that linear utility is equivalent to using expected value to predict choices: the options have different levels of risk, but the decision-maker is indifferent between them.

This idea was key in resolving the St Petersburg paradox: with a concave utility function, there would come a point where marginal utility plateaued, thereby anchoring the price that a rational player would be willing to pay for the coin flip game. Later developments on the expected utility hypothesis offered a second interpretation for the curvature of utility functions: a prediction of people's attitude vis-à-vis risk. Concavity led to risk-aversion (Fig. 1-3c), convexity led to risk-seeking behaviour (Fig. 1-3a), and a flat utility function was the equivalent of someone using expected values to make choices (i.e., without considering risk, Fig 1-3b). Beyond resolving the St Petersburg paradox and predicting people's attitude towards risk, however, expected utility's most important contribution is that it allowed for the comparison of different goods and outcomes using a *common currency* – this will later, as we will see, serve as the premise for identifying 'value' in the brain (Wu & Glimcher, 2018).

Today, Expected Utility Theory (hereafter EUT) stands as both the modern incarnation of Daniel Bernoulli's original hypothesis, and the nucleus of normative decision theory. In 1944, John von Neumann and Oskar Morgenstern proposed the four axioms of EUT, rules that individuals had to satisfy if their preferences were to follow the predictions of Bernoulli's expected utility hypothesis (when choices involved risky outcomes; von Neumann & Morgenstern, 1944). For an individual to be considered rational under the expected utility hypothesis, EUT states that their preferences must be: (i) complete, (ii) transitive, (iii) continuous, and (iv) independent of irrelevant factors. The first two axioms, completeness and transitivity, establish that decision-makers should be able to order their preferences. The continuity axiom, then, is necessary to establish a continuous scale upon which probabilities and outcomes are compared. The last axiom, that of independence, is necessary to establish that preferences keep their order regardless of other factors – such as the way choices are presented, the addition of decoy options (options that you would never pick anyways), or the scaling up or down of all options simultaneously. Where the first three axioms allow us to establish the general shape of a utility function, the independence axiom allows EUT to use said function in order to prescribe the choices one should make in different situations (in other words, it assumes preferences are static). Together, the axioms establish the conditions necessary for us to determine whether a decision-maker is acting 'as if' he is maximising expected utility. Von Neumann and Morgenstern, thus, mathematically defined rational choice theory in a way that could be used to rigorously study human decision-making.

It did not take long for EUT to become the dominant model of rational choice in economics (and decision theory, for that matter). In fact, most of neoclassical economics adopted von Neumann and Morgenstern's framework as an ideal mathematisation of rationality.

1.1.3 The limits of normative utility theory

As more economists and psychologists began using EUT to study decision-making, it became apparent that the source of its power, i.e., the axioms of rationality, would also prove to be its limitation. Most famously, in 1953, the French economist Maurice Allais presented a series of experiments in which decision-makers reliably violated EUT's independence assumption. Allais, who designed the two situations now known to as the *Allais paradox*, showed that people reliably violated EUT's predictions – particularly when it came to considering probabilities objectively (when computing the expected utility of risky outcomes).

The first violation of EUT exposed by Allais's Paradox came from adding a *common consequence* (or irrelevant alternative) to every option in a binary choice. Take the following two situations (adapted from Kahneman & Tversky, 1979):

> *"Situation 1:*
>
> *I offer you a choice between: option A, a 33% chance of getting £2500 (£0 otherwise); option B, a 34% chance of getting £2400. Do you prefer option A or option B?*
>
> *Situation 2:*
>
> *I offer you another choice between: option C, a 33% chance of getting £2500, 66% chance of getting £2400, and a 1% chance of £0; option D, £2400 for sure. Do you prefer option C or option D?"*

The decision is a hard one, but most people in situation 1 prefer option A. That is, they prefer the 33% chance of getting £2500 over the 34% chance of getting £2400. Given this result, EUT prescribes that, in situation 2, option C should be preferred to option D. This is not what Allais found. Instead, he found that people consistently behaved *irrationally* and picked option D. The simple addition of a common consequence led to different preferences in each situation – a scenario incompatible with EUT's axiom of independence. In numbers, we see that true preferences are incompatible with a single, consistent utility function:

$$[0.33 \times u(£2500)] > [0.34 \times u(£2400)] \quad (1\text{-}4)$$

$$[0.33 \times u(£2500)] + [0.66 \times u(£2400)] < [0.34 \times u(£2400)] + [0.66 \times u(£2400)] \quad (1\text{-}5)$$

Another violation of EUT highlighted by Allais is the one engendered by dividing or multiply decision variables by a *common ratio* (Allais, 1953; Camerer, 1989). An example of the common ratio effect, whereby the common ratio is an identical 75% reduction in the probability of getting each outcome, goes as follows (again, adapted from Kahneman & Tversky, 1979):

> *"Situation 1:*
>
> *Suppose I offer you a choice between option A, £3000, and option B, an 80% chance of getting £4000 (£0 otherwise). Which one would you prefer?*
>
> *Situation 2:*

> *If I offer you another choice between option C, a 25% chance of getting £3000 (£0 otherwise) and option D, a 20% chance of getting £4000 (£0 otherwise). Which one would you prefer?"*

For the first situation, most people pick option A, suggesting that most people are risk-averse, and that the expected utility of option A is higher than the expected utility of option B. From this result, and through the assumptions of EUT, it follows that in situation 2 people should pick option C. This, again, was shown to not be the case by Allais. Though each situation presented decision-makers with (i) the same monetary rewards, and (ii) scenarios where the second option occurred 80% of the time relative the first option's likelihood (100% vs 80%; 25% vs 20%), when presented with situation 2, people reversed their preferences and picked option D over option C.

$$[1.00 \times u(£3000)] > [0.80 \times u(£4000)] \tag{1-6}$$

$$[0.25 \times u(£3000)] < [0.20 \times u(£4000)] \tag{1-7}$$

Since we would need two different utility functions to capture the effects of both the common consequence and common ratio effects, Allais' Paradox represents a direct violation of von Neumann and Morgenstern's axiom of independence (Allais, 1953). Another, more popular interpretation, is that if we assume utility remains constant, what changes, is our evaluation of probabilities.

Until Allais, there had also been no clear challenge to the idea that uncertainty had to be treated objectively (unlike outcome magnitudes, which were transformed into utilities). Soon after Allais, however, Ellsberg's Paradox added fuel to the fire by demonstrating that uncertain outcomes with known probabilities (risky) are treated differently than uncertain outcomes with unknown probabilities (ambiguous). In fact, Ellsberg demonstrated that decision-makers reliably preferred risky outcomes to ambiguous ones (Ellsberg, 1961) – a behaviour that could, again, not be explained by von Neumann and Morgenstern's EUT.

The final nail in EUT's coffin came with the psychologists Daniel Kahneman and Amos Tversky's demonstration that, far from being independent of 'irrelevant factors', people's choices are inherently dependent on the way that choice outcomes are framed (Tversky & Kahneman, 1981, 1986). Specifically, they challenged the robustness of EUT's independence axiom with a series of simple experiments like the following (Tversky & Kahneman, 1981, p.453):

"Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

> *Situation 1:*
>
> *If Program A is adopted, 200 people will be saved.*

> *If Program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved.*
>
> *Situation 2:*
>
> *If Program C is adopted, 400 people will die.*
>
> *If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die."*

In these hypothetical scenarios, the outcomes for both the risky or riskless options are the same: 200 people survive, 400 people die. Still, when Kahneman and Tversky presented people with these choices, most preferred the safer program A in situation 1, but opted for the riskier program D in situation 2 – a direct violation of EUT's predictions (if A was preferred in situation 1, the equivalent program C should have also been preferred in situation 2). People reliably made decisions that were risk-averse when outcomes were presented as gains (as EUT predicts), but they became risk-seeking when outcomes were presented as losses – a direct reflection, or inversion, of their preference for risk.

What Kahneman and Tversky established with the above example is now known as the *framing effect*: how the way in which a choice is framed markedly changes the decision-maker's preference for risk. This effect has since been reproduced in a variety of situations (Henrich, 2004; McNeil et al., 1982), and in a variety of nonhuman decision-makers (we will cover this later in the chapter); it also forms the basis for decision theory's most successful 'descriptive' model of decision-making: *Prospect Theory* (PT).

1.2 Description

1.2.1 Prospect Theory

In 1979, Kahneman and Tversky proposed a novel, more descriptive model of decision-making that could account for the many violations of rational choice uncovered in previous decades. To do so, they strategically incorporated ideas from psychology into the framework of EUT, making for a more robust model of decision-making that nevertheless adhered to the principles of utility maximisation prized by normative economics (Figure 1-4). In a sense, they developed a new set of axioms and rules that made 'rational' the behaviours that EUT considered 'irrational' (Tversky & Kahneman, 1992; Volz & Gigerenzer, 2014) – they made utility maximisation descriptive. First, they addressed the framing effect (the main violation of independence) by integrating it directly into a decision-makers 'utility' function. PT replaced the wealth-dependent EUT utility function with a *value function* that interpreted outcomes as gains or losses relative to a decision-maker's reference-point. The value function maintained EUT's concavity (risk-aversion) for those outcomes considered gains, but became instead convex (risk-seeking) for outcomes considered losses (Fig. 1-4a). Second, the value function accounted for people's oversensitivity to losses (*loss aversion*) by creating a 'kink' at the value function's reference point. That

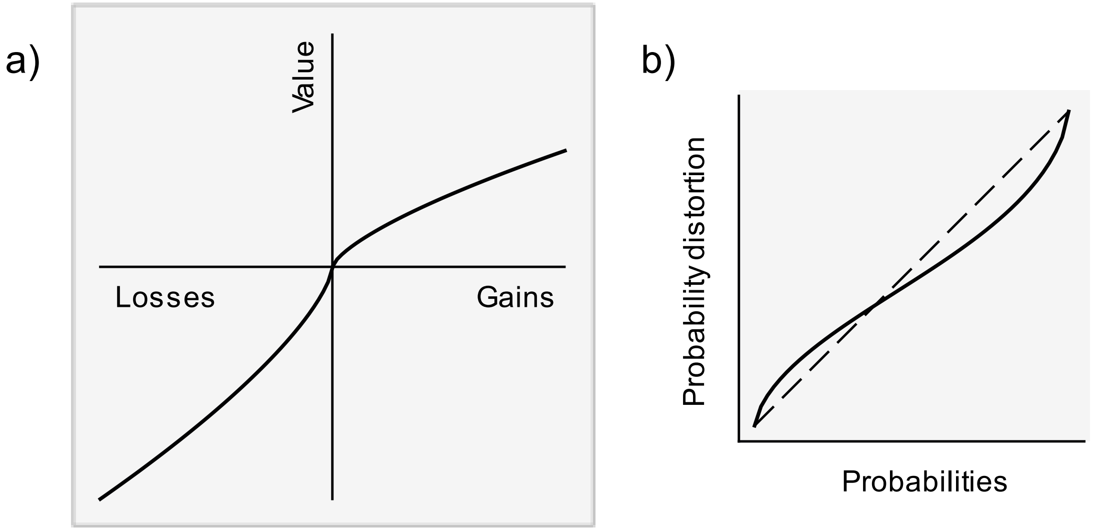


Figure 1-4 | Prospect theory's value and probability distortion functions. a) The value function's curvature reverses at the inflection point: it is concave for gains and convex for losses. The is also a 'kink' in the function at the reference, whereby losses are weighed more heavily than gain. b) People subjectively distort probabilities. Prospect theory's weighing function assumes that we overweigh low probabilities and underweight high ones.

is, the slope of the loss segment was more pronounced than that of the gains segment, such that a given loss required more than the equivalent gain to compensate for the decision-maker's extra aversion for losing. Third, probabilities were now also treated subjectively: outcome probabilities were weighed according to a distorted probability metric rather than the objective likelihood that they would occur. PT's weighing function (ω), or *probability distortion*, has since conventionally taken the shape of an inverted-S where low probabilities are overweighed, and high probabilities are underweighed (Fig. 1-4b). Integrating these three psychological concepts into the utility maximisation framework, the value of a prospect (a prospective outcome, or gamble) under PT was formalized as:

$$Value(x,p) \; = \sum \omega(p_n) \times v(x_n) \qquad (1\text{-}8)$$

where $\omega(p_n)$ is the distorted probability p_n that outcome x_n will occur, and $v(x_n)$ is the reference-dependent value of outcome x_n.

Through this simple but psychologically sound formulation, PT managed to account for many of the EUT violations published at the time (most famously, Allais' paradox). It also describes what has become known as the fourfold pattern of risk attitudes: that decision-makers are (i) risk-averse for risky gains if their probabilities are high, but (ii) risk-seeking for losses that have high probabilities (ii), and that they are (iii) risk-seeking for gains if their probabilities are low, but (iv) risk-averse for losses that have low probabilities (Scholten & Read, 2014; Tversky & Kahneman, 1992). Since it kept in line with the assumption of utility (value) maximisation, PT readily migrated from behavioural economics into the many other subfields of economics – becoming the leading model of choice in modern decision theory (Camerer, 1998; Camerer & Weber, 1992). For their development of PT, and, really, for pioneering the merger of psychology and economics that would become later come to be known as behavioural economics, Kahneman and Tversky were awarded the 2002 Nobel Prize in economics.

1.2.2 Heuristics and biases

Around the same time that they formalized PT, Kahneman and Tversky published a series of experimental results demonstrating that people did not always behave 'as-if' maximising some kind of internal 'utility' metric – an idea that had been previously advanced by the economist and psychologist Herbert Simon (Simon, 1966).

In Simon's words, the problem with rational choice and utility-maximisation (up to that point) had been the *'complete lack of evidence that, in actual human choice situations of any complexity, these computations [could] be, or in fact, [were] performed'* (Simon 1955). His solution, and perhaps Simon's most important contributions to decision theory, was the conceptualisation of *bounded rationality*: the idea that humans had a limit to their computational ability, and that this inherently constrained our decision-making abilities. In developing this idea, he suggested that humans, rather than calculating and comparing precise values and probabilities, often made us of mental shortcuts – rules of thumb he termed *heuristics* – to facilitate the decision-making process.

Kahneman and Tversky built on this concept through a series of experiments investigating people's use of heuristics to simplify complex decisions outside PT (Tversky & Kahneman, 1974). In their view, heuristics were used to reduce the level of complexity involved with calculating precise probabilities, and to simplify the 'rational' thought process used in predicting the value of our choices (Tversky & Kahneman, 1974). For example, a common heuristic they identified was that of *availability* (Tversky & Kahneman, 1973), wherein decision-makers assessed the likelihood of an event through the ease with which they could recall it (in other words, how available it is to them). Since the likelihood of something occurring is generally tied to our having experienced it, this is a good strategy in many settings. It can also, however, lead to suboptimal behaviour; particularly when dealing with events of low likelihood but high salience. For example, shark attacks rarely happen, and even more rarely lead to someone's death; yet, they have a very poignant effect on people, such that more people are more afraid of shark attacks than of drownings, of falling aeroplane debris (which in fact kill 30 times more people), or of falling coconuts (which kill twice as many people as shark; Barss, 1984).

In addition to the availability heuristic, Kahneman and Tversky also behaviours they termed the *Gambler's Fallacy* and the *Hot-Hand Bias*: opposing patterns of choice that are, nonetheless, both engendered by our tendency to perceive fake relationships between uncorrelated events (Falk & Konold, 1997; Nickerson, 2002; Rabin & Vayanos, 2010). The former, the gambler's fallacy, refers to a decision-maker's false belief that, in a sequence of unrelated events, those events that have not occurred for some time are more likely to occur – that things have to balance out (Tversky & Kahneman, 1971). In contrast, the hot-hand bias relates to the false belief that winning or losing comes in streaks (Gilovich et al., 1985).

Both phenomena have been tied to a heuristic that Tversky and Kahneman call *representativeness*: that we simplify the world through a belief that small samples (of events or variables) should be representative of the larger distributions from which they have been drawn (Kahneman & Tversky, 1972; Tversky & Kahneman, 1974). Using the earlier example of the coin flip game, one we understand as having a 50/50 chance of either a heads or tail outcome (because of the *Law of Large Numbers*[3]), the representativeness heuristic explains why most of us erroneously expect short bouts of play to mimic these underlying probabilities (we mistakenly believe in the *Law of Small Numbers*[4]). In the case of the gambler's fallacy, this easily translates to expecting heads after landing on tails multiple times. In the case

---

[3] When independently drawing from a distribution of outcomes, larger samples are more likely to represent the mean, or the expected value of said distribution.

[4] The incorrect belief that, when independently drawing from a distribution, even small samples should reflect the distribution expected value.

of the hot-hand effect, however, the reasoning is that representativeness leads people to interpret streaks as unrepresentative of (and therefore detached from) the randomisation process (Ayton & Fischer, 2004). In other words, the hot-hand effect makes us perceive streaks as special 'breaks' in the randomization process – ones we then expect will continue.

Though both phenomena have been described at length (for review, see J. Xu & Harvey, 2014b), there is still little consensus on these heuristics beyond their 'ad hoc' labels. In fact, behavioural economists still fundamentally disagree as to what might lead to one bias over the other, and about how these opposing biases can co-exist (Burns & Corpus, 2004; Miller & Sanjurjo, 2018; Oskarsson et al., 2009; J. Xu & Harvey, 2014b). A study by economists Rachel Croson and James Sundali (2005), looking at the decision-making of gamblers in casinos, perhaps best illustrates the problem. Focusing on the game of roulette, they found that gamblers (i) placed more bets after winning (i.e. hot-hand bias), but also (ii) bet significantly more against long streaks of numbers than with the long streaks (i.e. 5 repetitions or more). Both the hot-hand effect and the gambler's fallacy co-existed in the same people, but each heuristic targeted a different aspect of the same game – why? In a later study, the same authors found that these opposing patterns of behaviour also correlated in strength (Sundali & Croson, 2006), suggesting that, if not the representativeness heuristic, these two biases nonetheless shared some sort of underlying cognitive process (we will get back to this later).

By using the above heuristics to describe the behaviours that PT could not, Kahneman and Tversky led most of behavioural economics to view heuristics as unavoidable, imperfect substitutes to rational, utility-maximisation (Tversky & Kahneman, 1973). In doing so, their *heuristics-and-biases* research program popularised the idea that both rational and irrational choices could be explained - albeit separately (a distinction that still governs the descriptive approach today; cf. Fiedler & von Sydow, 2015; Gigerenzer, 2016).

1.2.3 Ecological rationality

In direct opposition to the work of Kahneman and Tversky, a branch of behavioural economics argues that the rational/irrational distinction is but an artificial one, and that the field might do better to understand the conditions under which 'irrational' heuristics are in fact optimal. In place of rational choice theory and the framework of utility maximisation, they define a new framework that stems directly from Herbert Simon's early criticism of 'as-if' theories and his later conceptualisation of bounded rationality. They call it *ecological rationality*, the evolutionary psychologist's take on behavioural economics (Hammerstein & Hagen, 2005; Hoffman et al., 1998).

Through an almost exclusive focus on the use of heuristics, ecological rationality argues that theories like PT offer no solutions to the failings of rational choice theory: they still assume near-infinite computational ability (i.e., unbounded rationality), and, in effect, only patch the failings of earlier rational choice models like EUT (Volz & Gigerenzer, 2014). In contrast, instead of a single, 'perfect' decision-making algorithm, the approach of ecological rationality is to view multiple, imperfect heuristics as the best course of action for decision-makers that have biological and cognitive constraints (Simon, 1990; Todd & Gigerenzer, 2000).

In line with this view, the psychologist (and architect of modern ecological rationality) Gerd Gigerenzer conceptualises the mind as an *adaptive toolbox:* a collection of simple but specialized heuristics that evolution has assembled into the different strategies we use to make choices (Gigerenzer & Selten, 2002). Unlike Kahneman and Tversky's heuristics, these should be specific enough to be formally model (i.e., not one-word labels; cf. Gigerenzer, 2010), yet also describe behaviour as accurately as - or better than - traditional models (Goldstein & Gigerenzer, 2002). Gigerenzer's heuristics are *fast-and-frugal* ones: simple algorithms that, by exploiting statistical regularities in our environment, facilitate choices when our knowledge is incomplete (Todd & Gigerenzer, 2000).

Take the aforementioned hot-hand bias (i.e. our tendency to incorrectly perceive positive correlations between the outcomes of random events). Although the rational choice school labels this bias as an 'irrational' application of the 'rational' law of large numbers, the approach of ecological rationality has been to try and identify situations where such a bias is in fact optimal. As a result, while traditional explanations for the hot-hand bias have focused on errors of representativeness (i.e. the law of small numbers), a separate set of studies has sought to formally understand the ecological realities that would favour a hot-hand bias (for review, see Kong, Granic, Lambert, & Teo, 2019). This rapidly growing body of work argues that the perception of streaks – far from irrational[5] – is in fact an optimal strategy: the hot-hand bias is significantly better than chance when dealing with uncertain or varying odds (e.g., Burns, 2004; Csapo, Avugos, Raab, & Bar-Eli, 2015; Hammack, Cooper, Flach, & Houpt, 2017; Miller & Sanjurjo, 2018b), and it presents no real disadvantages when dealing with true random processes (Scheibehenne, Wilke, & Todd, 2011). In contrast to that of the heuristics-and-biases view, the ecologically rational take on the hot-hand bias is thus that: "*if there is […] a pattern, expecting that particular pattern can be advantageous by providing an edge in predicting future events, and if there is no pattern, expecting one will not do worse than any other strategy*" (Scheibehenne et al., 2011, p. 327).

Supporting this interpretation, the hot-hand bias shifts in intensity as our brain ages (Castel et al., 2012), has been observed across different cultures (A. Wilke & Barrett, 2009), and is a heuristic we also share with other primates (Blanchard, Wilke, et al., 2014) – key indications that this bias, more than just a learned strategy, has an inherent biological basis (i.e. ripe for natural selection). Along that vein, a series

---

[5] The classic studies on the 'hot-hand' bias have been mostly carried out in relation to sports and sports betting, and assume that 'scoring' is generally uncorrelated (that is, players that score are not more likely to score again; see Gilovich, Vallone, & Tversky, 1985b; Kahneman, n.d.). Several studies have since, however, demonstrated that scoring is not 'as random' as initially thought, and that a belief in 'streaks' might not be as erroneous as they once seemed (for review, see Avugos, Köppen, Czienskowski, Raab, & Bar-Eli, 2013; or Bar-Eli, Avugos, & Raab, 2006)

of studies by the psychologists Andreas Wilke and Harold Clark Barrett (et al.) paint the hot-hand bias as a heuristic evolving to take advantage of the 'clumpiness' of our ancestral foraging environment (Scheibehenne et al., 2011; A. Wilke et al., 2009; A. Wilke & Barrett, 2009). Specifically, they suggest that looking for (and expecting) 'clumps' might be advantageous in a world where the distribution of plants, other animals, and water is often not random (A. Wilke et al., 2014). Evolving a heuristic, or bias, that piggy-backs on the clumped structure of our environment is thus a very ecologically rational thing to do. By taking advantage of statistical regularities in the environment, the hot-hand heuristic alleviates the need for more energetically expensive, 'rational' cognition (Mishra, 2014).

Through this *less-is-more* perspective, research on ecological rationality has identified 3 main rules, or building blocks, that generally compose heuristics: (i) *search rules* that specify how and where to look for choice-relevant information; (ii) *stopping rules* that specify when the search should end; and (iii) *decision rules* that specify how the final choice should be made (Todd & Gigerenzer, 2000). The specifics of these rules, and the ways in which they are tweaked, depend on the properties of the problem for which they have then been assembled. The *recognition heuristic*[6], for example, requires the decision-maker to search for information in memory (the search rule), to stop the search once something is recognized (the stopping rule), and to select the object that was recognized (the decision rule; see Gigerenzer & Goldstein, 2011). It is the more 'formal' analogue to Kahneman and Tversky's availability heuristic (unlike availability, however, it can be used precisely model choices). In a similar vein, the *priority heuristic* – a heuristic alternative to EUT or PT – posits that we *search* our options in terms of minimum gain, probability of minimum gain, and maximum gain (in that order); that we *stop* our examination if either the minimum gains differ by 1/10 or more of the maximum gain or if the probabilities differ by 1/10 or more of the probability scale; and that we *decide* on the option with the highest value for that variable that we were considering when we stopped (Brandstätter et al., 2006, 2008). With just 3 rules, and no mention of utility, the priority heuristic is in fact able to rival or outperform many of PT's predictions[7]: the preference reversals of the Allais paradox and the fourfold pattern of risk attitudes, to name a few.

Naturally, the fast-and-frugal view of heuristics clashes with the more predominant heuristics-and-biases one. Yet, for many non-economists, the heuristics of Gigerenzer's adaptive toolbox echoes the biological reality that is the selection of '*simple, repeated design patterns*' through human evolution

---

[6] The recognition heuristic, though analogous to Kahneman and Tverky's availability heuristic, differentiates itself via the fact that it is a clearly defined model of decision-making (not an informal, one-word label; Goldstein & Gigerenzer, 2002; Pachur, 2011). It also does not depend on the 'number of times' an event has been encountered, only the ability to recognize it(Pachur et al., 2011).

[7] It should be noted, nonetheless, that '*the priority heuristic models difficult decisions, not all decisions. It does not apply to pairs of gambles in which one gamble dominates the other one, and it also does not apply to "easy" problems in which the expected values are strikingly different*' (Brandstätter et al., 2006, p.8). In fact, in situations where decisions would be considered 'simple' the priority heuristic appears to fails quite remarkably (Rieger & Wang, 2008; but see (Brandstätter et al., 2008; Brandstätter & Gussmack, 2013).

(Adams, Watson, Pearson, & Platt, 2012), and is in keeping with the idea that the brain likely relies on a number of simple, repeated algorithms to build complex cognition (e.g., decision-making; McNamara & Houston, 2009; Pearson, Watson, & Platt, 2014). Since neural computations are energetically expensive, the use of simple algorithms that adapt to the structure of one's environment might represent the best cost-benefit tradeoff in terms of evolutionary investment (Hayden & Walton, 2014). What traditional economists view as suboptimal choices, then, likely reflects but a mismatch between (i) the evolutionary context in which our brain has assembled theses heuristics, and (ii) the artificial situations in which economic choices have been studied (Goldstein & Gigerenzer, 2002; Kacelnik, 2012). Put simply, fast-and-frugal heuristics are neither rational or irrational (in the classical sense); rather, they are ecologically rational if they lead to optimal choices in the situations for which they have evolved (Gigerenzer & Gaissmaier, 2011; Raab & Gigerenzer, 2015).

1.3 Prediction (or lack thereof)

Today, with ecological rationality still in its infancy, rational choice theory and its derivatives continue to dominate decision theory[8]. Following the trend set by Kahneman and Tversky, however, the last few decades have seen the behavioural economics inundated with newer models seeking to explain the failings of older ones (for reviews, see Camerer, 1989; Shafir & LeBoeuf, 2002). *Cumulative Prospect Theory*[9] rose to address the failures of PT (Tversky & Kahneman, 1992), new heuristics replaced old ones, and a collection of new theories and algorithms began to amass, each one describing specific choice biases in utmost detail (e.g., Köszegi & Rabin, 2006; Masatlioglu & Ok, 2005; Quiggin, 1982; Sugden, 2003). With no clear and unified framework to fall back on, the field has had no way of filtering this growing number of models (each one promising to outdo others in describing choices). In the words of Stephen Choi and Adam Pritchard (two eminent Wall Street regulators and law professors): "*Instead of a theory, behavioural economics [came to rely] on a hodgepodge of evidence showing the ineffectiveness of human decision-making in various circumstances"* (2003, p.10).

The solution to this problem, and the filter by which to consolidate decision theory, comes by way of the biological sciences…

---

[8] Just look at the way behavioural economics has been popularized in recent years. Books like Daniel Kahneman's *Thinking Fast and Slow* (2011) and Dan Ariely's *Predictibly Irrational* (2008), as well as the idea of *Nudging* in public policy (see Thaler & Sunstein, 2008), all stem from the rational approach. Specifically, from the 'divide' between rational and irrational thinking.

[9] Cumulative Prospect Theory (CPT), refines the original version of PT by addressing one of its original failures: that the sum of an option's distorted probabilities could, in effect, add up to more than 1 (creating all sorts of issues regarding the plausibility of its predictions). To address this issue, CPT incorporates the concept of rank-dependence, from rank-dependent EU (Quiggin, 1982, 1993), whereby it distorts the cumulative probabilities of options, not their objective ones (Fennema & Wakker, 1997; Tversky & Kahneman, 1992). It is this transform of PT that is used throughout this thesis.

1.3.1 The advent of neuroeconomics

In the 1980s, the field of neuroscience went through something of a revolution. The invention of novel imaging techniques suddenly allowed researchers that had previously been confined to the study of brain lesion (in animals and neurological patients) to explore a more direct link between the brain and behaviour. The study of decision-making through neurobiology had finally become a possibility, and many neuroscientists started using the choice models developed by behavioural economics to correlate the activity of different brain regions with choices and preferences.

The first successful study demonstrating a quantifiable link between brain and behaviour came in 1989 when neurophysiologists William Newsome and J. Anthony Movshon showed that the activity pattern of neurons in the monkey visual cortex correlated with the animals' choices (Glimcher & Fehr, 2014; Newsome et al., 1989). In a task where monkeys had to indicate the direction in which dots on a screen were moving, Newsom and Movshon found that the activity of individual neurons reliably correlated with the chosen direction. Then, three years later, the field saw the publication of the first three studies that used *functional magnetic resonance imaging* (fMRI[10]) to correlate whole-brain activity with behaviour (Bandettini et al., 1992; Kwong et al., 1992; Ogawa et al., 1992). The neuroscientific revolution had begun.

The development of fMRI in the early 1990s put neuroscience in an ideal position to explore the neural computations of human decision-making – something that economists had never before been able to do. Concurrently, behavioural economists started tooling with the idea that neuroscience could serve as the unifier they had been looking for: a tool that would identify the biological constraints that decision theory should place on models of choice. Before long, like-minded psychologists, neuroscientists, and economists began collaborating. Then, in 2005, a seminal paper by the behavioural economists Colin Camerer, George Lowenstein, and Drazen Prelec laid out the groundwork for what was to become the field of neuroeconomics (Camerer et al., 2005). In doing so, they outlined two unique and promising contributions of neuroscience to the study of decision-making:

(i) *Incremental, top-down contributions*: wherein neuroscience could refine or reform the 'as if' models of decision-making currently in use (which the authors argue "*have never been well supported empirically*").

(ii) *Radical, bottom-up contributions*: wherein the economic approach would be thoroughly re-evaluated, laying out a new neurobiological foundation to replace past economic theories.

---

[10] Unlike MRI, which produces a 'still' image of the brain, fMRI measures brain activity through changes in brain (blood) oxygenation levels over time. Since blood flow (and therefore oxygenation) is coupled with the needs of recently active neurons, fMRI signals allow us to pinpoint areas of the brain that were active at the time of the scan (Logothetis et al., 2001).

The goals were set, and the field of neuroeconomics was born. But while Camerer and his colleagues stressed that "*in the long run a more "radical' departure from current theory will become necessary"* (Camerer et al., 2005, p.55), the field's most influential work has stubbornly stuck to the incremental approach. The idea of a *neural common currency* (Glimcher et al., 2005; Levy & Glimcher, 2012; McNamara & Houston, 1986; Montague & Berns, 2002), a signal mimicking the economic concept of utility, has been particularly prevalent. An early study by the neurophysiologist Camillo Padoa-Schioppa and John A. Assad (2006), for example, identified the orbitofrontal cortex (OFC), a region of the prefrontal cortex that stands right above your eyeballs, as the one to represent the rational 'value' of our different choice options. By presenting monkeys with a choice between two juice options and simultaneous recording the activity of their OFC neurons, they found that (i) individual neurons encoded the value of either option separately, (ii) that this 'value' activity was indicative of choice, and (iii) that these values were independent of either option's position (i.e. the motor response needed to select it). A follow-up study, reiterating the OFC's role in computing value, showed that within a task, the neuronal representations of value suggested transitivity: that is, the value of each option was independent of the other options available to the monkey – signals were absolute, not relative (Padoa-Schioppa & Assad, 2008). Studies like these have since guided a volley of others into identifying different brain regions with activity patterns that correlate with EUT's utility or PT's value, their axiomatic components, or the broader ideas of rational choice theory (O'Doherty, 2014; O'Neill & Schultz, 2018; Pastor-Bernier et al., 2017; Stauffer et al., 2014). Conversely, with most of neuroeconomics following the assumptions rational choice, there has been but a handful of studies looking at the neurobiological plausibility of ecological rationality (Hunt & Hayden, 2017; Khader et al., 2011; Kotz & Dorp, 2010)

Though arguably one-sided, neuroeconomics' focus has permitted the identification of brain networks involved in the behavioural manifestations of PT's framing effect (De Martino et al., 2006; Zheng et al., 2010), loss aversion (Knutson et al., 2008; Tom et al., 2007) and probability distortion (Hsu et al., 2009; Tobler et al., 2008). And while the main contributions of neuroeconomics remain incremental, the study of neural signals – particularly their 'noisiness', or *stochasticity*[11] – has been central to the development of the novel choice models put forth in recent years (see Hunt & Hayden, 2017; Louie, Glimcher, & Webb, 2015; Rustichini & Padoa-Schioppa, 2015; this will be covered in more depth in the **interlogue** and **postface** chapters).

Though the incremental approach has delivered on its refinement promise, neuroeconomics still has a long way to go if it is to radically change our understanding of economic decision-making (Bossaerts & Murawski, 2015). Specifically, rather than using 'as-if' models to locate and quantify choice-related activity in the brain, neuroeconomics needs to start questioning the neurobiological validity of the models it uses (i.e. act as a filter; cf. Volz & Gigerenzer, 2014). Many of these models, for instance, show a

---

[11] The inherent randomness with which neurons process information and communicate

remarkable lack of predictive power. An entire book, '*Risky curves: On the empirical failure of expected utility'*, has been written on the topic (D. Friedman et al., 2014). How can we expect to accurately map out (and indeed predict) brain function if none of the models we use accurately predict out-of-sample choices? A paradigm shift is on the horizon for neuroeconomics: one that will hopefully take it from predicting biology with economic models, to predicting economic choices with biology (something the field can uniquely do).

1.3.2 Comparative psychology and primate economics

Whereas neuroeconomists typically see the use of animals as a window into the human brain (i.e. neural systems from which we can record neural activity[12]), the field of comparative psychology offers a different take on the matter. If decision theory is to truly understand human decision-making, they argue that the study of animal decision-making is, in of itself, a worthwhile endeavour. Animals with whom we share an evolutionary history offer us a glimpse into the past – a way to understand the evolutionary forces that have shaped our decision-making. In studying them, comparative psychology poses important questions for decision theory: (i) can we use the models of human decision-making to successfully describe the behaviour of animals that share in our neurobiology (other primates, in particular); and (ii) if not, why?

The study of animal choices, and indeed the comparative approach, is perhaps best illustrated by a famous series of experiments on capuchin behaviour. In 2006, Chen, Lakshminaryanan, and Santos made headlines[13] following their publication of a study that showed capuchin monkeys consistently selecting food rewards that were presented as gains over identical food rewards presented as losses (just like the framing effect PT describes in humans). Following this, in 2011 the same group went on to demonstrate that capuchin monkeys were risk-averse for choices presented as gains, but risk-seeking when choices were presented as losses (just like humans; Lakshminarayanan, Chen, & Santos, 2011). Their procedure was elegantly simple: two experimenters presented capuchins with two sets of apple pieces. The number of pieces would then vary in a known or unknown way before the monkey received whichever set they had chosen. In the gains frame, the safe experimenter always added an additional piece of apple to the starting one; the risky experimenter sometimes added two, sometimes added nothing. The loss frame was similar, but instead of starting with a single piece of apple and adding more, the experimenters consistently removed one piece from an initial three, or probabilistically removed two. As expected, the monkeys were risk-averse for choices framed as gains, picking the 'safe'

---

[12] It should be noted that neuroscientists resort to animal experiments only when there is no alternative, or when it is impractical or unethical to use specific treatments/methods on humans (things like electrode recordings, selective lesioning, or genetic manipulations). Several neuroscientists described this best in a 2018 report: *"studies, using invasive methods, complement those in human volunteers and patients, using non-invasive methods such as brain imaging and computational modelling"* (Mitchell et al., 2018).

[13] I would recommend to any reader to watch Laurie Santos' TED talk on the subject, titled 'A monkey economy as irrational as ours': https://www.ted.com/talks/laurie_santos?language=en

experimenter significantly more than the 'risky' one; and they were risk-seeking for choices framed as losses – even though the outcomes were the same.

Until then, though we knew animals of all kind had a lot of experience choosing between primary biological rewards (e.g., food, water, sex), it had been less obvious that they should behave like humans (especially 'rational' humans) when participating in similar economic experiments. The work of Yale's Comparative Cognition Laboratory changed all of that: it showed that human models could be used to describe the behaviour of primates (the effect being replicated in many different animals, including chimpanzees and bonobos; see Krupenye, Rosati, & Hare, 2015).

The application of PT to the study of primate behaviour has its limitations, however, and it becomes apparent when we attempt to generalize decision strategies across species: risk attitudes are not consistent across tasks, across species, and even within single species. Humans and capuchins might be largely risk-averse for choices involving gains, but many primates instead demonstrate risk-seeking preferences. Macaques, for example, are generally risk-seeking in neuroeconomic experiments (McCoy & Platt, 2005; O'Neill & Schultz, 2010; Smith et al., 2017; Stauffer et al., 2015; E. R. Xu & Kralik, 2014), but a handful of studies also demonstrate risk-averse behaviour (Genest et al., 2016; Yamada et al., 2013). Bonobos and Chimpanzees, two of the most related primate species, also show widely divergent patterns of risk attitude: chimpanzees appear to be more risk-seeking, while bonobos exhibit human-like risk aversion (Heilbronner, Rosati, Stevens, Hare, & Hauser, 2008; Krupenye et al., 2015). Conversely, a study by Haun et al. (2011), in which the risk attitude of all four great apes was compared, found that they were all risk-seeking to some degree, while another study found that orangutans were instead risk-averse (Pelé et al., 2014). This last one, wherein orangutans, capuchins, and macaques had to decide between (i) keeping a cup of cookies they had been given or (ii) taking a gamble and trading with the experimenter for a different one (out of six randomly selected ones from a pile they could observe), was particularly revealing. They found that orangutans and capuchins were risk-averse, but that long-tail macaques were not (just like rhesus macaques).

This diversity in risk attitudes in primates, often inconsistent within single species, is confusing at every level of decision theory. Rational choice cannot explain it, and ecological rationality alone does not provide a clear biological basis that evolution could act on. Still, where we used to think that differences in decision-making represented fundamental differences in neurobiology, comparative psychology has shown that simple, fundamental biases, exacerbated by task design, are often the ones to blame (Heilbronner, 2017; Volz & Gigerenzer, 2014). Macaques overweight gambles that they have previously won (Blanchard, Wilke, et al., 2014), humans and monkeys become more risk-seeking when playing for relatively small rewards (Heilbronner & Hayden, 2013; B. J. Weber & Chapman, 2005), and both monkeys and humans underweight rewards when their probabilities that are stated overtly compared with those where the probabilities have to be learned (Heilbronner & Hayden, 2016; Hertwig

& Erev, 2009; B. J. Weber & Chapman, 2005). In parallel, a new wave of neuroeconomics has also begun to highlight the effect of past choices and contextual statistics in shaping neural activity and the behaviour that ensues (e.g., Diederen, Spencer, Vestergaard, Fletcher, & Schultz, 2016; Webb, Glimcher, & Louie, 2014; Zimmermann, Glimcher, & Louie, 2018). It is, in fact, the effects of contextual statistics on decision-making that this thesis later explores.

1.3.3 The promise of neuroethology

With the merger of behavioural economics and neuroscience giving rise to the field of neuroeconomics, a similar merger, this time between *ethology* (the study of natural animal behaviour) and neuroscience, produced what we now know as neuroethology (Konishi, 2010). A field focused on natural behaviour, it differentiates itself from neuroeconomics in that it considers the effects of evolutionary pressure as fundamental to the study of the nervous systems. To truly understand the brain, it argues that neither behaviour nor the circuits that generate it can be taken out of their evolutionary context (a view not too distant from that of ecological rationality). The evolutionary origins of a species, their development, as well as the anatomical and physiological limitations that biology engenders, are all central to the neuroethological perspective.

The modern incarnation of the field, originating in the 80s, has only recently garnered the interest of those that study decision-making (Adams et al., 2012; Watson & Platt, 2008). Nevertheless, the neuroethological approach is likely the key to a more radical rework of decision theory – one more in-line with what Camerer et al. had intended (2005). Unlike the neuroeconomists[14] who gaze into the brain using the ‘as-if’ models of decision theory, neuroethologists build models that consider: (i) the choice problems animals confront, (ii) the algorithms that could be used to solve these problems, and (iii) most importantly, whether the brain’s architecture can actually support said algorithms (Pearson et al., 2014). In doing so, neuroethology offers a bottom-up take on the neurobiology of decision-making and brings a much needed second opinion to an area where theories have previously gone unchallenged.

Along these lines, one of neuroethology’s clearest, and most robust finding is that the algorithms we use to make choices are flexible; that, biologically, *homo sapiens* is closer to ecological rationality’s *homo heuristics* than rational choice’s *homo economicus* (cf. Gigerenzer & Brighton, 2009; Kalenscher & Wingerden, 2013; Levitt & Lis, 2008). Ethology has been arguing for years that behavioural flexibility is what has made us primates so successful (Jones, 2005). Likewise, studies that look at the neural basis of our more ‘natural’ decisions are starting to highlight the dynamic nature our brain’s choice algorithms (e.g., Blanchard & Hayden, 2014; Hayden & Walton, 2014; Kolling, Wittmann, & Rushworth, 2014; Kvitsiani et al., 2013). Research on primate foraging behaviour, for example, is

---

[14] For simplicity’s sake, I use the term ‘neuroeconomics’ to capture those neuroscientific studies that describe and predict neurobiological variables using economic theory; and ‘neuroethology’ as those studies that instead build theories from the neurobiology. Though the dichotomy is, in reality, much blurrier, imposing a clear-cut distinction is useful for this discussion and to understand the power of those studies that blur the line.

particularly insightful. A study on the by the neuroscientists Benjamin Hayden, John Pearson, and Michael Platt (2011), had rhesus macaques make choices between two options that emulated the choices a monkey could make when foraging in a 'patchy' environment (i.e. one where water and food is found in 'clumps'). Termed the patch-leaving problem, monkeys were faced with the decision to either stay and get a reward (juice with diminishing returns) or leave the patch and reset their 'staying' earnings. Every time they picked the stay option they got a little bit less reward, but every time they chose to leave they faced a travel delay (time) associated with the 'search' for the next patch. While monkeys were choosing, Hayden et al. found that the activity in their cingulate cortex[15] mimicked an *accumulation-to-threshold* process: the activity of individual neurons slowly built up as the animals 'depleted' the patch and got closer to their 'leave' decision; once activity reached a specific threshold, the monkey would choose to leave (it largely chose to stay, otherwise). Interestingly, though the activity threshold remained constant, the increments with which neurons 'built-up' activity changed depending on the time delay involved with leaving. The brain seemed to be comparing a relatively stable representation of the value of staying, with a relatively dynamic value of leaving (the flexibility needed in a changing environment; Hayden, 2018).

Accumulation-to-threshold models, of which the most popular is *drift-diffusion* (for review, Ratcliff, Smith, Brown, & McKoon, 2016), have long been a staple of perceptual decision-making research (a branch of neuroscience that focuses on problems of recognition, or detection of motion direction; e.g., Summerfield & De Lange, 2014). Mimicking the way neurons in our sensory systems make choices, these models predict that individual groups of neurons, representing different but specific options, ramp up their firing activity as perceptual evidence accumulates. The first group of neurons to reach a specific activity threshold (or segment of the population) leads to that group's respective option being selected by the system. On a psychophysical level, when applied to value-based decision-making, these models highlight that the relative difference in the activity of neuronal populations is what determines choices. Perhaps surprisingly, accumulation models also suggest that the absolute 'value' of our options might significantly influence the time it takes for us to 'make up our mind' (Busemeyer & Townsend, 1993; Hunt et al., 2012; Webb, 2018).

Another characteristic of cortical circuits, discovered again by the study of perception (e.g., Britten & Heuer, 1999; Carandini, Heeger, & Movshon, 1997; Heeger, 1992*;* but see also, Howard & Kahnt, 2017; Louie, Grattan, & Glimcher, 2011; Strait, Blanchard, & Hayden, 2014), is that individual neurons encode information relative to the activity of their neighbours (for review, see Carandini & Heeger, 2012). The activity of neurons, in other words, is *normalized* in a context-dependent manner – they become more sensitive to high magnitude stimuli when such signals are abundant, but they become less sensitive to high magnitude stimuli if low magnitude signals become the norm (Summerfield & Tsetsos,

---

[15] The dorsal anterior cingulate cortex, more precisely.

2015). According to the neurobiologists Simon Laughlin, this is a classic case of *efficient-coding:* because of the physical limits to the information they can represent[16], the most efficient thing for neurons to do is match their input-output function to the distribution of inputs they might expect (Fig. 1-5a). In Laughlin's words, "*The strategy of matching a neuron's input-output function to the expected distribution of signals [...] ensures that the maximum amount of information is transferred between elements*" (Laughlin, 1981, p.912). Just like the heuristics of ecological rationality, however, efficiency sometimes produces odd– if informative – results. In the visual system, for example, normalization is behind various optical illusions (for example, see Figure 1-5b). Similarly, in value systems, normalization can lead to some of rational choice theory's most famous violations (Louie et al., 2013) and likely serves as the basis for reference-dependent preferences, as described by PT (LoFaro et al., 2014; Louie et al., 2014).

A final neurobiological reality worth mentioning, and one made most famous by my own PhD Supervisor, is the near ubiquity of *reinforcement learning* signals in the brain (e.g., Doll, Simon, & Daw, 2012; Vickery, Chun, & Lee, 2011, for review, see Schultz, 2015). In 1997, the neurophysiologist Wolfram Schultz and his colleagues, the computational neuroscientists Peter Dayan and Read Montague, published a defining study on the unique role of dopamine in predicting and updating the value of rewards. Where midbrain dopamine neurons had been previously known to predict appetitive stimuli (think of dopamine's implication in drug addiction), Schultz and his colleagues demonstrated that they also signalled what machine learning experts called a *reward prediction error*. By measuring dopaminergic activity in the brain of monkeys during learning, the team showed that dopamine neurons emitted "*a positive signal (increased spike production) if an appetitive event [was] better than predicted, no signal (no change in spike production) if an appetitive event occur[ed] as predicted, and a negative signal (decreased spike production) if an appetitive event [was] worse than predicted*" (1997, 1594). Schultz et al.'s study demonstrated that dopamine's role was twofold: first, it encodes information about the subjective value and context of the rewards we expect to get; and second, it highlights any mismatch between reality and expectations in a way that promotes learning throughout the brain (Bayer & Glimcher, 2005; Schultz, 2016; Stauffer et al., 2014). By modulating synaptic plasticity, this second dopamine signal it is theorized to reinforces rewarding behaviours while discouraging unrewarded ones (Daw & Tobler, 2013). In computational theory, we call this phenomenon reinforcement learning, whereby a system or algorithm 'learns' from past actions in order to better predict and optimize future choices (Sutton & Barto, 2018).

The prevalence and recurrence of these basic features of the brain point to the very real possibility that, rather than using complex and inflexible models, decision theory might be better served by sets of simple but adaptive algorithms (Kuchibhotla et al., 2017; Yoo & Hayden, 2018).

---

[16] The limits of the cortex' neural code is generally estimated to be between 0 and ~200 action potentials a second.

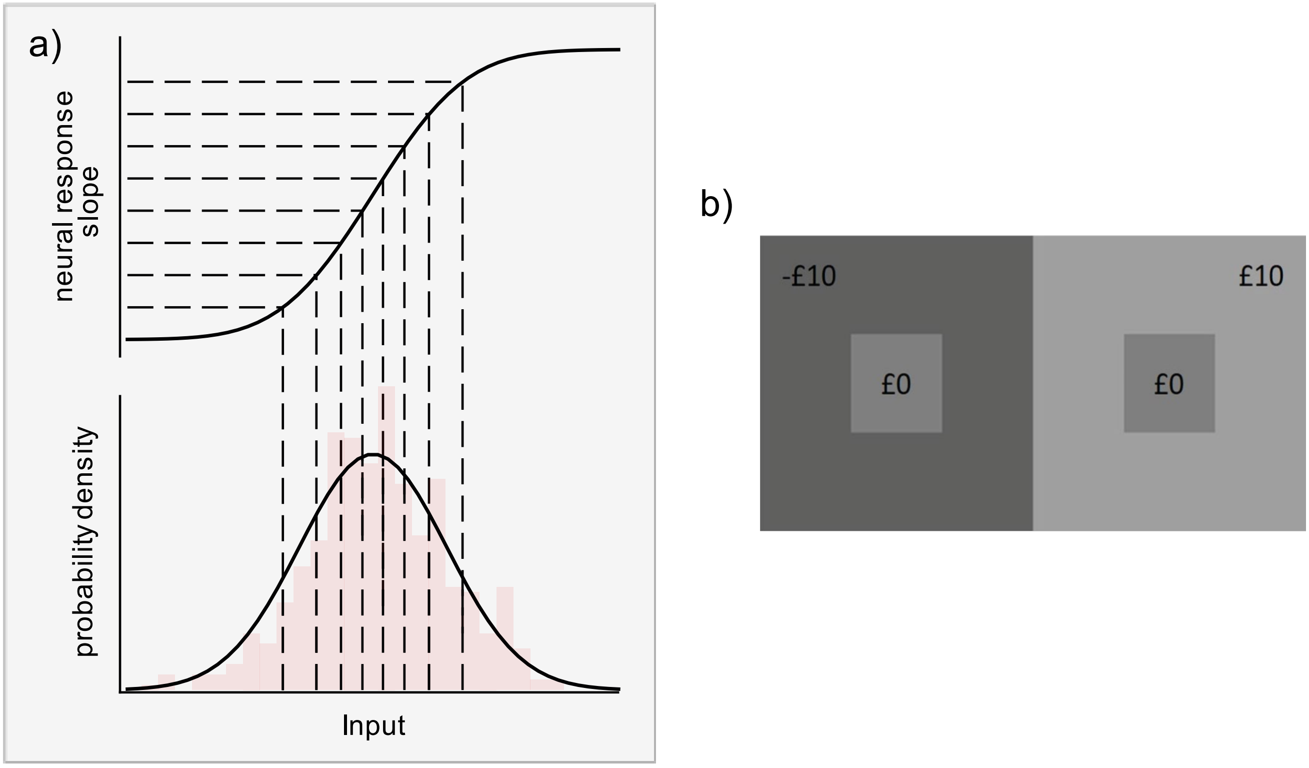


Figure 1-5 | Relative coding as a defining feature of the brain. a) Neurons efficiently encode their inputs. To optimize both their energy use and their limited coding capacity (i.e. max/min firing rates) neurons tune their processing (the input-output function; top) to best represent the distribution of their inputs (bottom). By tuning their response function, neurons allocate more perceptual resources to those stimuli that occur more frequently; in turn, neurons become more precise in encoding differences between these stimuli (Adapted from *Laughlin, 1981*). b) Optical illusion from spatial context. The central squares on either side of this figure are the same colour: a medium shade of grey. The way our brain interprets this colour, however, depends on the context in which we find it. If the background is dark, then relative the background the central square looks lighter. If the background is light, then relative to the background the central square looks darker. According to neuroscientists, the same kind of process governs the way we evaluate rewards. 0£ looks better if our expectation was a loss of 10£; conversely, 0£ looks worse if we expected to win 10£ (Adapted from *Louie & Glimcher, 2012).*

Just as ecological rationality had suggested, we are now finding that the brain's decision algorithms are simple enough to be efficient, but flexible enough to be useful in a wide array of situations (Kalenscher & Wingerden, 2013). Neurons encode values in a context-dependent way, they learn from experience, and they finetune their firing rates to match the realities of our environment – particularly in those areas, like the orbitofrontal cortex, previously though to hold our neural representations of utility (Diederen et al., 2016; Shunsuke Kobayashi et al., 2010; Padoa-Schioppa, 2009; Rustichini et al., 2017; Zimmermann et al., 2018). Why not use these properties of neurons to – if not fully redefine decision models – at least edit existing ones in a way that makes them more plausible. There is some evidence to suggest that reinforcement learning could induce hot-hand biases, for instance (Abrahamyan et al., 2016; Neiman & Loewenstein, 2011; Worthy & Todd Maddox, 2014), as well as in shaping the way we value repeated, uncertain decisions (e.g., G. Barron & Erev, 2003; Lejarraga & Gonzalez, 2011; Weiss-Cohen, Konstantinidis, Speekenbrink, & Harvey, 2018; Wulff, Mergenthaler-Canseco, & Hertwig, 2018). As an argument to incorporating flexible algorithms into the formal study of risky, decision-making, **Chapters 2 and 3** will investigate the effects of adaptation on choice, while **Chapter 4** focuses on the situation described just above.

1.4 Bringing decision theory full-circle

As you may have come to realize, though the many subfields of decision theory complement each other by focusing on different questions (recall Fig. 1-1), the way they 'knit' their ideas into a cohesive whole ignores a very important question: how does all of it tie-in together?

Broadly speaking, the different fields and questions reviewed here look at decision-making from either a top-down or a bottom-up perspective. The former, the top-down approach, (i) looks at choices, (ii) suggests algorithms that can describe these choices, and (iii) searches for neural representations of these algorithms in the brain. Economists build models from the choices people make, psychologist refine them, and neuroeconomists use them. The bottom-up approach, in contrast, should do the reverse: (i) start from the basic biological circuits we find in the brains, (ii) use these to describe the algorithms that the brain can support, and (iii) filter for the algorithms that match real, observable choices. The problem with decision theory is that, while the two perspectives could ideally form a circular framework (where top-down and bottom-up inform and build on each other; see Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017), the two are wholly unbalanced. Besides, even if decision theory started building from the ground up, in humans we are limited in that we cannot precisely measure single-cell activity (non-invasive imagine precludes a truly bottom-up perspective). Animal research, on the other hand, presents itself as a clear candidate for a truly circular theory of choice,

1.4.1 Animal experiments as uniquely interdisciplinary

Though the research presented in this chapter has, until now, paints animal research mostly as a means to validate human findings, the study of animal decision-making provides much more, and is better represented by the following divisions in its approach:

1. The *comparative approach*, looks at animals other than humans in an effort to highlight continuities and discontinuities in choices across species (e.g., Brosnan, Wilson, & Beran, 2012; Krupenye, Rosati, & Hare, 2015; Paglieri et al., 2014; Pelé, Broihanne, Thierry, Call, & Dufour, 2014; Prétôt, Bshary, & Brosnan, 2016; Uhlrich, Micheletta, Thierry, Dufour, & Pelé, 2010; Watzek & Brosnan, 2018).
2. The *neurophysiological approach,* uses animals to explore the computations of single neurons using experimental methods that we could otherwise not use in humans, e.g., single-cell recordings, lesion studies, and optogenetics (e.g., Boyden et al., 2016; McCoy & Platt, 2005; Murray et al., 2014; Padoa-Schioppa, 2009; Padoa-Schioppa & Assad, 2006; Platt & Glimcher, 1999; Schultz, 1998).
3. The more 'traditional' economic *modelling approach*, uses animal experiments to validate and refine economic models in optimally controlled environments (Ferrari-Toniolo et al., 2019; Genest et al., 2016; Hayden, 2018), and to propose new models for humans that build on animal behaviour (foraging theory, for example; Houston, Fawcett, Mallpress, & McNamara, 2014; Mobbs, Trimmer, Blumstein, & Dayan, 2018; Stevens, 2010).

Each of these approaches offers a unique take on the study of decision-making, exploring fundamental questions that would otherwise be left unexplored in human experiments. These also tie in well with one of the most popular, and circular frameworks for studying and understanding cognitive processes: David Marr's levels of information processing, or the idea that to fully understand a cognitive system one must first understand it via three distinct but complementary levels of analysis (Marr & Ullman, 1982). We need to study (i) the 'goals' of the decision-making system, (ii) the cognitive software or algorithms that can support these goals, and (iii) the neural circuits that run the cognitive software (Figure 1-6).

In line with David Marr's levels of analysis, comparative research ties in well with investigating the computation, or goal of the decision-making system (what we want to achieve). Behavioural ecologists, ethologists, and psychologists compare the choices of human and nonhuman animals to explore biases and choice strategies in terms of evolution. This provides valuable insight as to which characteristics of human decision-making are commonly shared and which ones are lost as we move away from homo sapiens; we can then use this information to generate hypotheses about the evolutionary origins of specific decision strategies (for review, see Santos & Platt, 2013; Santos & Rosati, 2015). Understanding how simple decision patterns evolved – and how they were maintained – goes a long way in helping us understand and model human decision-making strategies. In a sense, this approach is the closest thing we have to travelling through time and observing evolution's effects on the reliability of economic models.

The second approach, wherein animals are used to directly probe the neurons involved in decision-making, ties in with Marr's implementation level and fills a unique niche in decision-making research. While human neuroimaging (fMRI, MEG, EEG) has mostly worked from the top-down, using economic models to identify regional brain activity we think correlates with decision-making, single-cell recordings have been instrumental in understanding the basic computations supported by neurons (Cox & Kable, 2014; Huk, 2005; Schultz, 1998; Shadlen & Newsome, 2001).

The final branch of animal economics rounds off Marr's levels of analysis with a focus on testing and developing new algorithms that maximally predict the choices of humans and other animals. Here, behavioural economists, computational neuroscientists, and psychologists focus on improving the robustness and predictive ability of various models through animal experiments. They then use these findings to better inform models of human choice behaviour. The premise for this approach is that if we can find or modify existing algorithms to reliably describe choices across multiple species, we can be more confident in their ability to predict our own (Sugden et al., 2006).

The separation between these approaches parallels the questions asked by human research: comparative studies seek to discover what the differences are between the choices of different species; neurophysiologists try to see how choices arise in the brain, and modellers look at why we might observe different patterns of choice in different situations. As we have seen in the previous section, these questions, if isolated, fall into separate fields of decision theory; the power of animal research is that it allows for the merger and integration of these approaches into one cohesive whole.

We can, within a single animal or species, study all three of Marr's levels, such that we can fully integrate the top-down and bottom-up perspectives on decision-making (Pearson et al., 2014). Comparative studies have often paved the way for neuroscientists and economists to integrate evolutionary comparisons into their own research. Using normative and descriptive economic theories, comparative studies can model the behavioural differences observed across species. Neuroscientists can then use these differences to investigate the neurobiological features that support them. Likewise, if the goal is to understand how ecological rationality shapes choices, ecologists can explore the individual evolutionary niches of different species to understand how these niches have shaped the decision strategies of animals. Where the bottom-up perspective fails in human research, however, the animal-oriented field of neuroethology (the study of how nervous systems generate natural behaviour in animals) works precisely to pinpoint the cognitive algorithms supported by neurobiological findings (Konishi, 2010; Pearson et al., 2014; Tinbergen, 1963) – something more akin to the radical approach promised by neuroeconomics (Camerer et al., 2005). In line with the idea of ecological rationality, incorporating

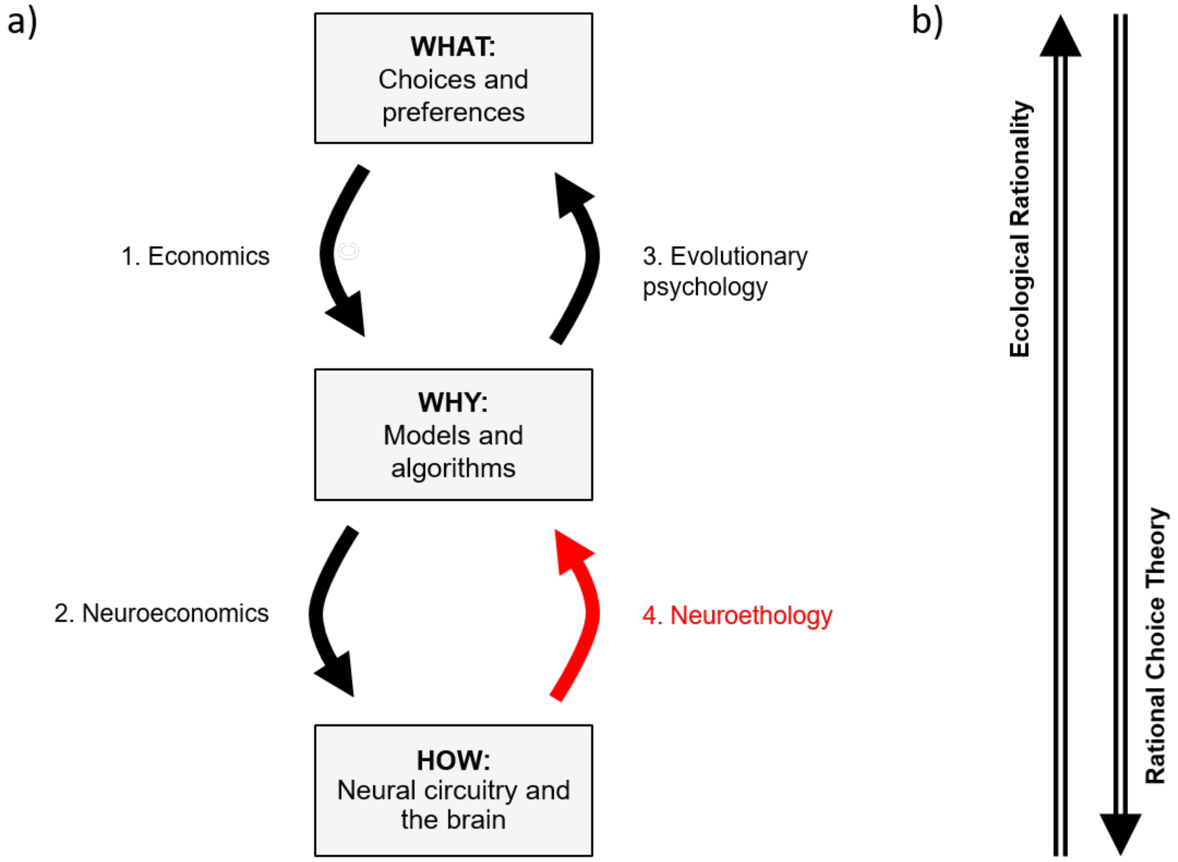


Figure 1-6 | The missing link in decision-theory: neuroethology. a) Re-illustrating Figure 1-1 with the appropriate decision-theory disciplines serving as the link between the questions that we can ask. What questions tie-in with Marr's computational level, why question tie in with the algorithm (or software), and how questions relate to biology (the hardware). Neuroethology has been mostly ignored in decision-theory up until now, but it forms the crucial link in unifying the different frameworks presented here. b) Top-down and bottom-up perspectives in decision theory. The former refers to using the choices we observe to ultimately explain the neural circuitry that supports decision-making. The latter looks at neural circuitry directly and suggests models that best predict the choices we observe. The two perspectives need to work together to form a circulate and complete analysis of decision-making.

ethology and neuroethology into decision theory provides a new, alternative set of hypotheses that can help reconcile the 'rational' and 'irrational' behaviours established by the top-down approach in decision theory.

1.4.2 Reframing choices from the bottom up

This thesis stands at the confluent of the ideas presented in this chapter. In fact, the aim of my entire PhD has been to try and reconcile the findings of decision theory's bottom-up and top-down frameworks.

We have known for some time now that that activity all over the macaque brain adapts to the statistical properties of the task at hand (Carandini & Heeger, 2012; Laughlin, 1981; Louie et al., 2015; Tremblay & Schultz, 1999) – yet, most top-down research in economics, psychology, and indeed neuroeconomics still relies on models that hold preferences as fixed. Likewise, economics has decades of research that quantifies systematically the effects of choice context in shaping people and macaque's preferences (Barberis, 2012; Blanchard, Wilke, et al., 2014; Blanchard, Wolfe, et al., 2014; Heilbronner, 2017; Hertwig et al., 2004; Kahneman & Tversky, 1979; Machina, 1987) – reproducing these effects should be a goal of the bottom-up approach.

The work that follows builds from both neuroscience's current understanding of the biological processes that support decision-making, and on the decades of economic research that have come to define decision theory. With this in mind, the following chapters attempt to answer one, unifying question: what if, rather than relying on stable preferences, we could predict economic choices from preferences that adapt. The logic goes as followed: if decision-making relies on the activity of neurons, and the activity of neurons adapt to changes in our environment, then preferences should adapt to changes in the environment. The trick is then to refine algorithms so that they explain economic choices, mimic the activity of neurons, and fulfil the requirements of ecological rationality… Easy!

Methods Interlogue

Econometrics and discrete choice modelling

Where this thesis' introduction focused particularly on the many 'irrational' choices that people make, the problem of inconsistency is another issue one has to acknowledge when trying to study real, repeated behaviour. In a study in 1989 by the American economists Colin Camerer, for example, 31.6% of experimental subjects were shown to reverse their preferences when repeatedly presented with the same binary choices; the economists Starmer and Sugden, also in 1989, published similar findings where 26.5% of participants changed preferences. What started as a couple of peculiar studies has now turned into the norm: people are inconsistent, and economists need models that can describe these inconsistencies (for a review, see Stott, 2006).

To interpret these results, economists began toying with the idea that *deterministic* choice models like Expected Utility Theory (EUT) and Prospect Theory (PT) - which fail to account for the variability across choices – might be better served by a bit of imprecision (Cartwright & Festinger, 1943; Machina, 1985). New models began to accommodate for the noise and errors that seemed so inherent to the decision-making process (Becker et al., 1963; Block & Marschak, 1959; Hey, 2005), and new axioms were developed that allowed for utility-maximising under assumptions of inconsistency (Abdellaoui, 2000; Gonzalez & Wu, 1999a; McFadden. Daniel, 2014; D. L. McFadden, 2005). The solution that eventually came about, *stochastic choice models*, is now one of the most common ways by which econometrist model choices: rather than predicting choices in a yes or no manner, these models evaluate the probability of a decision-maker making a specific choice (Harless & Camerer, 1994; Loomes & Sugden, 1995; McFadden, 2005).

One of the building blocks of stochastic choice models is the Luce choice axiom (Pfanzagl & Luce, 1959). A product of 1950s psychology, Luce's axiom provides the mathematical foundation for the synthesis that eventually became stochastic choice modelling. Put simply, it states that choices are probabilistic and that the probability of choosing a given option from one set of alternatives is directly proportional to the probability of choosing the same option from a larger set of alternatives (Luce, 1977). Mathematically, it takes the following form:

$$\mathrm{P}(x_i) = \frac{v(\mathrm{x_i})}{\sum_{\mathrm{j}} v(x_j)} \qquad (\text{ i-1 })$$

whereby $\mathrm{P}(x_i)$, the probability of choosing a given option x_i is dependent on the strength of x_i's representation ($v(x_i)$ as the internal representation, utility, or subjective value) relative to that of all other options (x_j) in a choice set. Under Luce's formulation, the probability of choosing a specific option can be directly related to a metric similar to that of utility (Luce, 1958; Machina, 1985). In other words, it represents a probabilistic interpretation of the axiom of independence so crucial to EUT. Most contemporary stochastic choice models (also knowns as discrete choice models) rest on this early mathematical assumption.

After Luce next step in accounting for inconsistencies within rational choice theory takes the form of Jacob Marschak's Random Utility Maximisation model (or RUM; 1960). Using Thurston's Law of comparative judgement, whereby objective stimuli are said to be perceived with a certain degree of noise, Jacob Marschak postulated that the utility of an outcome, at any given time, was the result of a true utility value to which noise was added: a direct integration of stochasticity, a 'randomness' component, into the framework of utility maximisation (Becker et al., 1963; Marschak, 1960). Under Marschak's assumption that decision-makers sought to maximize this *random utility*, the probability that a given option x_i would be chosen was given by:

$$\mathrm{P}(x_i) = \mathrm{P}[U(x_i) \geq U(x_j)] \text{ for all } i \neq j \qquad (\text{ i-2 })$$

$$\mathrm{P}(x_i) = \mathrm{P}[u(x_i) + \varepsilon_i \geq u(x_j) + \varepsilon_j] \text{ for all } i \neq j \qquad (\text{ i-3 })$$

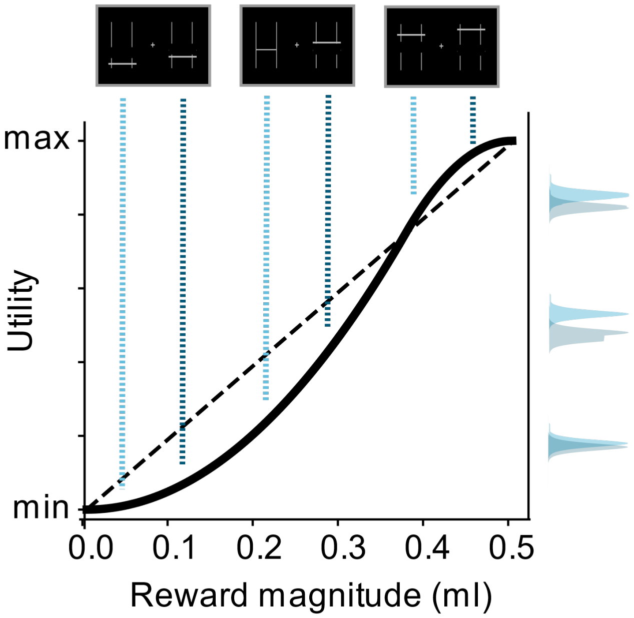
$$\mathrm{P}(x_i) = \mathrm{P}[u(x_i) - u(x_j) \geq \varepsilon_j - \varepsilon_i] \text{ for all } i \neq j \qquad (\text{ i-4 })$$

Where the probability of choosing x_i is the probability that the utility of x_i, with error ε_i, being larger or equal to x_j, with error ε_j. Importantly, the error terms ε are independent, but come from equivariant distributions.

The integration of this RUM framework with the Luce choice model allowed the economist Daniel McFadden, in 1968, to computationally model real choice behaviour (published later, McFadden, 1974a, 1974b). Reinterpreting Luce's equation into what we now refer to as a multinomial logit model (or logistic regression), McFadden showed that:

$$\mathrm{P}(x_i) = \frac{\mathrm{e}^{\mathrm{v}(x_i)}}{\sum_{\mathrm{j}} \mathrm{e}^{\mathrm{v}(x_j)}} \qquad (\text{ i-5 })$$

Where $\mathrm{P}(x_i)$ could now be directly estimated as a parametric function – the slope of the logit function capturing the noise in the decision-maker's choices. To this day, McFadden's logit model allows experimental economists to model choice algorithms using real experimental data, and to explore the validity of utility-maximisation models with realistic assumptions about our decision-making abilities. The work presented in this thesis stands as one such example, and relies almost entirely on the principles of discrete choice described in this **interlogue**.



Interlogue figure | The link between utility and stochastic choice. Assuming utilities are noisy and can be reinterpreted as equivariant distributions centred around a 'true' utility function, the stochasticity of choices can be directly interpreted as an overlap on the cognitive representations of 'random utilities'. Steeper utility functions should lead to less overlap in random utilities and more precise choices; flatter utility functions will lead to more overlap and therefore more stochastic choices. Represented above are three sets of choices between 2 safe (certain) reward. The certainty with which monkeys (or humans, for that matter) make choices will depend on the overlap between the random utility distributions of the rewards involved.

Chapter 2

Partial adaptation to reward range induces shifts in rhesus macaque risk preferences

2.1 Introduction

Every day we make choices between outcomes that vary widely, sometimes on the order of magnitudes. In a single morning, we can go from choosing between outfits, to choosing to visit our favourite cafe, to comparing the costs of a train or plane journey for our next holiday destination. Yet, despite the complexity of representing all of these situations, we manage - with a relatively limited brain - to mentalise and indeed optimise the majority of our choices.

Prospect Theory (PT), the dominant model in behavioural economics, posits that we optimize our decisions by calculating the value of our choices relative to a reference-point (Kahneman & Tversky, 1979; Tversky & Kahneman, 1986). That is, rather than objectively evaluating the outcome of our choices, we perceive our options as gains or losses depending on what we are expecting: if the outcomes are better than our reference, we treat them as gains; if they are worse, we treat them as losses. Mathematically, PT represents this behaviour with an S-shaped value (or utility) function where the subjective value of gains and losses is given by concave and convex parts of the function, respectively. This has important behavioural consequences, particularly for risky-decision-making, as this normative (utility) framework predicts that people's tendency to make risk-averse decisions depends on their perception of outcomes as being gains or losses.

While the idea of reference-dependence has been readily adopted by modern decision theory (Rabin, 2000; Wakker, 2010), economists are still unclear about how reference points form (Barberis, 2012). In prospect theory (PT), Kahneman and Tversky abstractly define reference-points as exogenous from the decisions being made. That is, the reference point is not directly explained by PT and can be shaped by "*aspirations, expectations, norms, and social comparisons"* (A. Tversky & Kahneman, 1991, p.157). Alternatively, some economists suggest that reference points are but an epiphenomenon of the way in which our mind adapts to the statistics of the task at hand (Delquié & Cillo, 2006; Köszegi & Rabin, 2006; Sugden, 2003) - a framework more in line with the findings that, far from being restricted to human reasoning, reference-dependence is a homogeneous feature of animal decision-making (Santos & Rosati, 2015) and the brain (Carandini & Heeger, 2012; Louie et al., 2013; Padoa-Schioppa, 2009;

Tremblay & Schultz, 1999). Along these lines, one particularly interesting proposal from the epiphenomenon framework, is that of range-dependent utility, or RDU (a play on reference-dependent utility; see Kontek & Lewandowski, 2018). Inspired by psychology's *range-frequency theory* (Parducci, 1965, 2012) and neurobiology's *efficient-coding hypothesis (*recall Laughlin's work in **Chapter 1)**, RDU suggests that decision-makers evaluate the value of their options relative to not one, but two reference points: the minimum and maximum rewards available in any given scenario. In this view, what PT identifies as a reference-point could be nothing more than the product of a range-adapting utility function: the point at a sigmoidal curve inflects from convex to concave (mimicking a neuron's tuning curve; Carandini & Heeger, 2012; Webster, Werner, & Field, 2005).

Because studies on reference-dependence generally focus on identifying a unique reference-point (Baillon et al., 2015), or on describing behaviours under specific reference predictions (Allen et al., 2016; Crawford & Meng, 2011; Wenner, 2015), there is, as of yet, no way of corroborating or contradicting the previous hypotheses on the emergence of reference-points. The few studies that consider shifts in preferences generally do so in a single-range, local context: they document reference-point changes following the wins or losses of risky gambles (Arkes et al., 2008, 2010; Shi et al., 2015); never the impact that changes expectations have on decision-making. Concurrently, little is known about the impact of a task's structure on preferences, nor how different reward statistics might translate to reference-points.

Animal experiments offer a unique avenue whereby we can collect extensive, robust, and controlled behavioural data in a variety of scenarios, and explore the formation of reference-points when exogenous factors have minimal impact (i.e. no contribution of language or higher numerical ability). To that effect, I investigated how the range of rewards experienced in a binary choice task shaped the preferences of rhesus macaques – a species that displays many, if not most, of the fundamental choice patterns humans display (Heilbronner & Hayden, 2013, 2016; Stauffer et al., 2015). I presented macaques with several sets of risky choice options where the range of possible reward magnitudes remained stable for weeks at a time, then suddenly shifted new ranges (higher/lower magnitudes or wider/narrower spread). On each testing day, I fit the monkeys' choices with S-shaped utility functions that could explain both risk-seeking and risk-averse choices (Genest et al., 2016; Stauffer et al., 2014), I then looked at how the monkey's risk preferences changed as a function of the reward ranges they experienced. I found that, while utilities stayed relatively put for periods where a single reward range was experienced, the monkeys consistently shifted their preferences when a novel reward range was introduced. What the monkeys did not do, however, was rely solely on the current reward statistics to form their reference; in fact, while preferences adapted easily to novel reward magnitudes, the monkeys' reference also kept track of the lowest rewards they had experienced over the course of the preceding weeks – even if these now fell outside of possible rewards. From our findings, I suggest that preferences are unlikely to reflect a single reference or expectation in a given reward range, but rather, that preferences follow the expectation of what monkeys think might happen given the knowledge they have accumulated over time.

2.2 Methods

2.2.1 Macaque monkeys as economic decision-makers

The current study was conducted with non-human primates, specifically, three male rhesus macaques (Macaca mulatta) weighing 11.2, 15.3, and 13.2 kg (monkeys A, B and C, respectively). All macaques in the study were born in captivity, at the Medical Research Council's 'Centre for Macaques' in the UK. The monkeys were pair-housed for most of the experiment; monkeys B and C shared an enclosure. The monkeys ranged in age from 5 to 8 years old, and all subjects had previous experience with the visual stimuli and experimental setup (Ferrari-Toniolo et al., 2019). The experimental protocol was assessed and approved by the Home Office of the United Kingdom.

Rhesus macaques are the world's most common old-world primate. In fact, they are the second most widely distributed primate species after humans (Timmins et al., 2008). Their geographical range extends from Afghanistan to China, passing through India and most of south-east Asia – and they thrive near human settlements. Rhesus macaques are an incredibly adaptable species, and while our evolutionary paths split nearly 25 million years ago (Kumar & Hedges, 1998; Rogers & Gibbs, 2014), we share with them a remarkable array of physical and psychological traits. For this reason, rhesus monkeys are the most commonplace species of non-human primate found in scientific research (Capitanio & Emborg, 2008). This has led to the development of a rich literature reproducing human economic choices in rhesus macaques, and one I build from in this study. Most relevant here is that rhesus macaques behaviour can be successfully described using PT (Farashahi et al., 2018; Ferrari-Toniolo et al., 2019; Genest et al., 2016; Stauffer et al., 2015). In addition, macaque experiments allow us to control the pre- and post-experimental environments in ways not possible for human studies – I can ensure that experimental variables are independent of rewards and choices made outside of the experiment (K. M. Chen et al., 2006). For this study, the delivery and range of rewards experienced were unique to the experimental setup. Unlike humans with monetary pay-outs, the monkeys experienced nothing comparable outside of the laboratory.

2.2.2 Behavioural task and training

Three male monkeys used a left-right joystick (Biotronix Workshop, University of Cambridge) to make choices between reward-predicting stimuli presented on a computer screen[17]. After each choice, the monkeys received their chosen reward in the form of a specific blackcurrant juice quantity delivered probabilistically (matching the probabilities indicated by each stimulus).

---

[17] All experimental protocols were assessed and approved by the Home Office of the United Kingdom.

The animals were presented with a simple visual scheme of one or two horizontal lines positioned between a frame of two vertical lines that depicted reward options that varied both in terms of magnitude (liquid quantities, m) and in terms of probability (the likelihood of a reward being delivered, p). Reward magnitudes were represented by the vertical position of the horizontal lines on the screen, while reward probabilities were represented by the relative width of the horizontal lines within their vertical line frame (Fig. 2-1a). As a result, safe (riskless) options were represented by singular full-width horizontal lines that touched both sides of their vertical frame, whilst multiple horizontal lines within the vertical frame signalled gambles, i.e., multiple 'risky' rewards.

The monkeys were trained to associate these two-dimensional visual stimuli with blackcurrant juice rewards over the course of >10,000 single-outcome, imperative trials. In these trials, a single reward option was presented on either the left or right side of the screen. To obtain the cued reward, the animals were required to select the side on which the reward was presented. Following imperative training, all experimental data were gathered within a binary choice paradigm where the monkeys had to choose one of two reward options presented simultaneously. One option was always a gamble; the other was always a safe, guaranteed reward. Every choice trial began with a white cross at the centre of a black screen, followed by the appearance cursor. To initiate a trial the monkeys had to, using the joystick, move the cursor to the centre cross and hold it there 0.5-1s. After a successful central hold, two reward options appeared left and right of the central cross (Fig. 2-1a). The monkeys had 3s to convey their decision by moving the joystick to the selected side and holding their choice for 0.1s to 0.2s - the unselected option would then disappear. The selected option lingered on the screen for 1s after reward delivery to strengthen any stimulus-reward associations with visual feedback. A variable intertrial period of 1–2 s (blank screen) preceded the next trial onset. Errors were defined as unsuccessful central hold, side selection hold, or trials where no choices were made, and resulted in a 6s timeout for the animals, after which the trial would be repeated.

All reward options were repeated on the left and right sides of the computer screen, alternating pseudorandomly to control for any side preference. Both the joystick position and task event times were sampled and stored at 1 kHz on a Windows 7 computer running custom MATLAB software (The MathWorks, 2015a; Psychtoolbox version 3.0.11), and all further analyses were done using custom Python code (Python 3.7.3, Scipy 1.2.1, Oliphant, 2007). Over the course of 63, 43 and 57 sessions an average of 259 ± 154 (mean ± STD) trials, 317 ± 118 trials, and 131 ± 75 trials were collected for monkeys A, B and C, respectively. Crucially, monkeys received the reward they selected after each and every one of these trials. This ensured that they experienced the rewards they selected with minimal and constant delay, and contrasts with human studies where only a randomly selected subset of trials are rewarded at the end of experimental sessions. Delivering rewards after every trial also allowed us to capture preferences that were contingent on experiences unique to the task - similar reward delivery and ranges were not experienced in the housing environment.

2.2.3 *Measuring preferences in range-specific conditions*

To examine the degree at which preferences are shaped by a task's available rewards, binary choice data were collected from choices between reward options affixed to different reward ranges (Fig. 2-1b). Three range conditions were defined in terms of their mean reward magnitude and the range of possible options: 1) low narrow-range, where tested magnitudes were generally set between 0 ml and 0.5 ml; 2) high narrow-range, with magnitudes between 0.5 ml and 1.0 ml; and 3) full range, with magnitudes between 0 ml and 1.0 ml (0.1 to 1.3 ml for monkey C). Importantly, every reward outcome (no matter the range) was repeated the same amount of times for each session – that meant that every reward was technically equiprobable (they were presented and delivered according to a flat distribution) The laboratory environment allowed us to set range conditions and keep them fixed for days at a time; measuring the effects of reward range over weeks rather than blocks (Fig. 2-1c).

Choice preferences were measured in each range by presenting individual monkeys with a series of gamble-safe choices from which Von Neumann–Morgenstern type utilities were estimated (utilities under the assumption of linear probability distortion[18]). Specifically, the utilities associated with the juice magnitudes presented in each experimental range were estimated using the multi-step fractile method (Genest et al., 2016; Lak et al., 2014; Machina, 1987).

For every range tested, a utility of a 0% and 100% was assigned to the lowest and highest reward magnitudes possible, respectively. The use of utility percentages, rather than 0-1 units, was adopted here to simplify comparisons between utility functions that fundamentally differ both in terms of scaling and range. From this, and because probabilities were assumed to be treated objectively, it followed that the equiprobable gamble formed by the minimum and maximum rewards in a range would be worth 50% utility (50% = [0.5 * 0%] + [0.5 * 100%]). Each testing session began by presenting this 50% utility gamble (m1 = min reward, p1 = 50%, m2 = max reward, p2 = 50%) against safe options of varying magnitudes (steps of 0.05 ml). The ratio of safe choices for each gamble-safe pair was recorded, and a logistic function was fit across all the measured ratios for the gamble:

$$P_{ChooseSafe} = 1/(1 + e^{-\left(\frac{SafeReward_{ml} - x_0}{\sigma}\right)}) \tag{2-1}$$

The probability of choosing the safe option ($P_{ChooseSafe}$) depended on the safe reward's magnitude ($SafeReward_{ml}$) and two free parameters: x_0 indicated the x-axis position of the curve's inflexion point, and σ, the temperature parameter. The curve's inflexion point, set at $P_{ChooseSafe} = 0.5$, also

---

[18] PT predicts that, since every gamble had a probability of p=0.5, every gamble would be underweighed or overweighed by the same quantity vis-à-vis safe rewards. In that case, ignoring probability distortion in the analysis skews the s-shaped utility function – but it does so in a way that is constant across all measurements. Since changes in range magnitude should, theoretically, only affect utility – we chose to ignore probability distortions and focus instead on changes in utility across reward ranges.

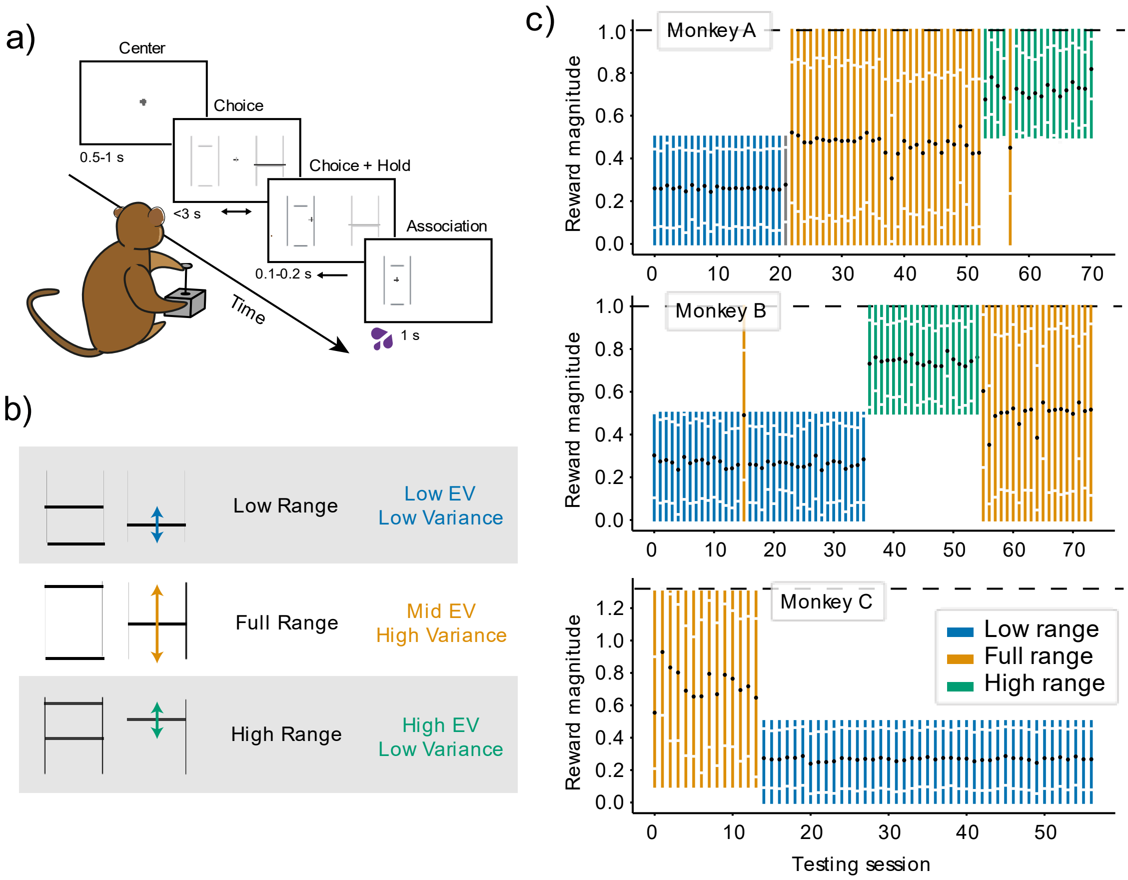


Figure 2-1| Experimental design and timescale. a) Binary choice task. The monkeys chose one of two gambles with a left-right motion joystick. They received the blackcurrant juice reward associated with the chosen stimuli after each trial: the reward's magnitude and probability of delivery were signalled by the vertical position and width of a horizontal line as set between two vertical ones. Times, in seconds, indicate the duration of each of the task's main events. b) Experimental reward ranges. Choices were made in one of three experimental reward ranges. In the low range, choice options had juice magnitudes set between 0 ml and 0.5 ml during preference elicitation sequences. The high range involved juice magnitudes set between 0.5 ml and 1.0 ml during preference elicitation sequences (unique to monkey A and B). The full range was set between 0 ml and 1.0 ml for monkeys A and B and set between 0.1 ml and 1.3 ml for monkey C. c) Monkeys experienced specific reward ranges for consecutive days. Vertical lines represent the daily experimental session, in their tested order; the height of these lines signals the reward range tested (blue, low range; yellow, full range; green high range). Black dots indicate the mean magnitude of all rewards experienced on the day, the white dots represent the standard deviation on the mean.

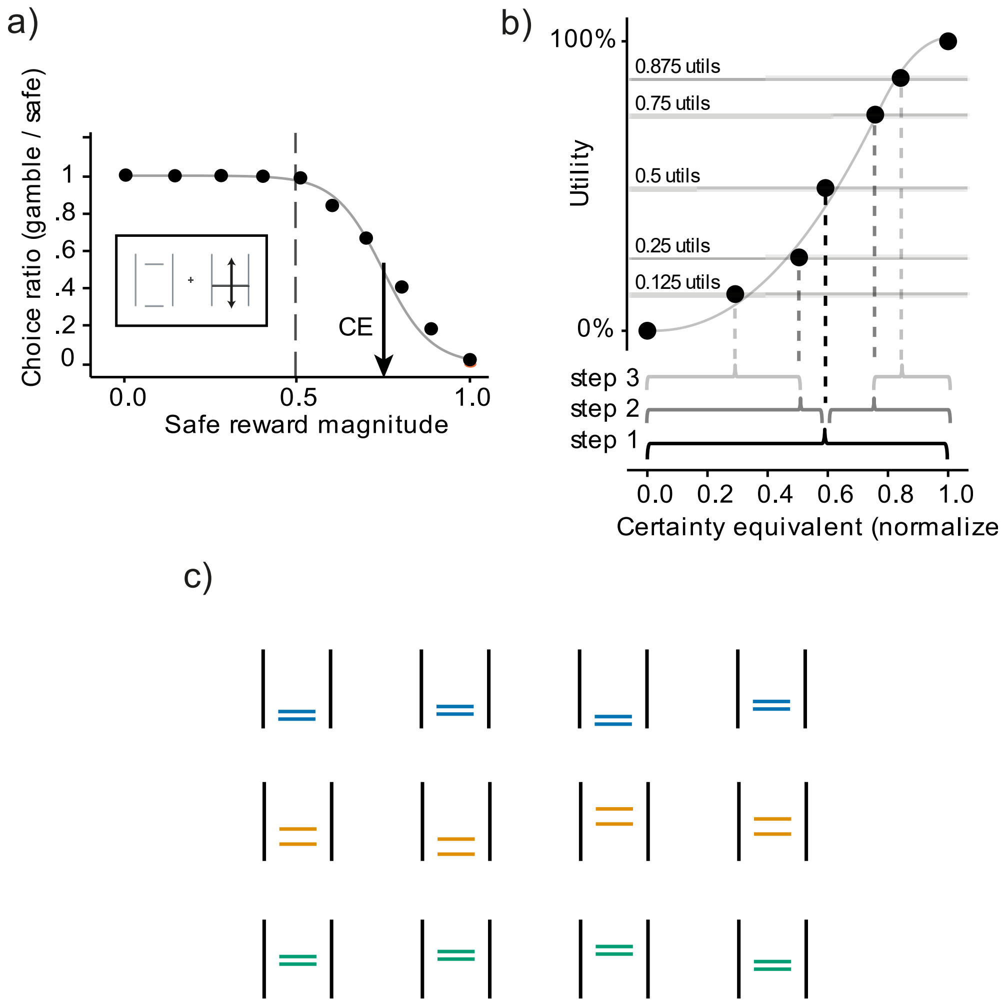


Figure 2-2 | Measuring certainty equivalents and utilities. a) Basic choice behaviour and estimation of certainty equivalents. Monkeys chose between a safe reward and a gamble on each trial. The safe rewards alternated pseudorandom on every trial – never going above or below the highest and lowest magnitudes tested in the daily reward range. Each point is a measure of choice ratio: the monkey's probability of choosing the gamble option over various safe rewards. We fit psychometric softmax functions (Eq. 2-1) to these choice ratios, separately for each day, and recorded the certainty equivalent (CE) of individual gambles as the safe magnitude for which the probability of either choice would be 0.5 (black arrow). The dashed vertical line indicates the expected value (EV) of the gamble represented in the box. b) Estimation of utility using the stepwise, fractile method. In step 1, the monkeys were presented with an equivariant gamble comprised of the maximum and minimum magnitudes in the tested reward range. the CE of the gamble was estimated and assigned a utility of 50%. In step 2, two new equivariant gambles were defined from the CE elicited in step 1. The CEs of these gambles were elicited and assigned a utility of 25% and 75%. Two more gambles are defined in step 3, from the CEs elicited in step 2. Their CEs were then assigned a utility of 12.5% and 87.5%. Parametric utility functions, anchored at 0 and 1, were fitted on these utility estimates (see *methods*). c) Equivariant, equiprobable gambles presented in out-of-sample validation sequences. Sets of four gambles, unique to each reward range, were used to validate the risk attitudes predicted by the fractile-derived utilities. The CEs of these gambles were measured (see panel a) and the difference between CEs and the specific gambles' EVs signalled the monkeys' risk attitudes: if the difference was positive, the animals were risk-seeking, if the difference was negative, the animals were risk-averse.

captured the gamble's certainty equivalent (CE): the safe reward magnitude for which the monkey would be choice indifferent. Taking the gamble's utility value as 50%, the gamble's CE represented a direct functional mapping of reward magnitudes to this utility (Fig. 2-2).

After the first CE had been elicited, two new equiprobable gambles were created: one between the lowest reward magnitude and the first CE, the other between the first CE and the highest reward magnitude. This sectioned the utility range in two equal halves, and the earlier CE elicitation procedure was repeated on the gambles defined from these two halves. Gamble-safe pairings for the elicitation of both gambles' CEs were interweaved in a single block, and utilities of 75% (75% = [0.5 * 50%] + [0.5 * 100%]) and 25% (25% = [0.5 * 0%] + [0.5 * 50%]) were assigned to the higher and lower gambles' estimated CE, respectively. Finally, a final set of gambles was defined from these CEs – and their individual CEs were mapped to utilities of 12.5% and 87.5%. Only sequences that contained a minimum of three different choice pairs (repeated at least 4 times) were used in the elicitation of CEs, and only the fractile sequences where at least 3 utility values could reliably be estimated were used in further analyses. The CEs assigned to each utility level, in each reward range, were compare via two-way ANOVA.

2.2.4 Parametric estimation of utilities from aggregate and single choices

Parametric utility curves were fit onto the CE-Utility data to capture and predict an animal's choice preferences over the entire range of rewards. These utility curves served as a direct signal of the animals' risk attitude over the tested reward range: if the fitted utilities were convex (i.e., increasingly curving upwards) the monkeys had demonstrated risk-seeking behaviour; if the curves were instead concave (i.e., gradually flattening), the monkeys had demonstrated risk aversion. Several parametric utility models were compared to ensure the most reliable utility predictions; the best-fitting functions would then be used for all further analyses. In accordance with the assumptions of the fractile method, each of these functions had to be anchored at 0% to 100% on the y-axis — and we normalized the CEs on which they were fit to be between 0 and 1. Finally, because CEs, not utilities, were the measured data (i.e., the error was relative to the x-axis), orthogonal distance regression was used to fit each and every function (Boggs & Rogers, 2012). I fit two 1-parameter functions ($U_{1\text{-Power}}$, $U_{1\text{-Tversky}}$),

$$U_{1-power}(m) \;=\; m^{\alpha} \qquad (\,2\text{-}2\,)^{19}$$

$$U_{1-Tversky}(m) \;=\; \frac{m^{\alpha}}{(m^{\alpha}+(1-m)^{\alpha})^{1/\alpha}} \qquad (\,2\text{-}3\,)^{20}$$

In each of these functions, m stands for the juice magnitudes (in ml) of a given reward outcome.

[19] $U_{1\text{-Power}}$: the α-parameter is the power parameter of the function (if $\alpha < 1$ utility is convex, if $\alpha > 1$ utility is concave).

[20] $U_{1\text{-Tversky}}$: the α-parameter is the temperature parameter of the function (if $\alpha > 1$ utility is S-shaped, if $\alpha < 1$ utility is inverse S-shaped).

two 2-parameter functions ($U_{2\text{-Prelec}}$, $U_{2\text{-SCDF}}$),

$$U_{2-Prelec}(m) \; = e^{-\beta\times(-\,ln(m))^{\alpha}} \qquad (\ 2\text{-}4\ )^{21}$$

$$U_{2-SCDF}(m) \; = \begin{cases} \beta \times \left(\frac{m}{\kappa}\right)^{1/\alpha}, for\ 0 \leq \ m \leq \ \beta \\ 1 - \ (1-\beta) \times \left(\frac{1-m}{1-\beta}\right)^{1/\alpha}, for\ \beta < m \leq 1 \end{cases} \qquad (\ 2\text{-}5\ )^{22}$$

And one 3-parameter function ($U_{3\text{-Power}}$)

$$U_{3-power}(m) \; = \begin{cases} (m-\gamma)^{\alpha}, for\ m \geq \ \gamma \\ -\beta \times (\gamma - m)^{\alpha}, for\ m \ < \ \gamma \end{cases} \qquad (\ 2\text{-}6\ )^{23}$$

Sets of daily Bayesian Information Criterions (BIC) were then calculated from the orthogonal residuals of each fitted model ($BIC_{RSS} \; = \; n \times ln\left(\frac{residuals}{n}\right) + (k \times ln(n))$). I selected the best fitting function using a one-way Friedman test followed by pairwise Wilcoxon signed-rank tests (Bonferroni-Holm corrected) and compared the estimated parameters specific to each reward range using a one-way MANOVA.

Since the fractile method relied on stepwise, chained measurements (where later metrics depend on earlier ones), utility functions were also estimated using a discrete choice model (DCM) fitted to single trials for comparison. By fitting a model on individual choices rather than aggregate CE sequences, I avoided the propagation of estimation errors from earlier steps onto the next and therefore reduced estimation biases for individual utility functions (Abdellaoui, 2000).

As is commonly done (McFadden, 2001; Stott, 2006), the likelihood that monkeys would choose the left option over the right one, given a set noise level and side bias, was modelled using a logit function:

$$P_{chooseLeft} \; = \frac{1}{(1 + e^{-\lambda(EU_{Left} - EU_{Right} \, - \, \theta)}} \qquad (\ 2\text{-}7\ )$$

The probability of choosing the left option was, therefore, in the DCM, a function of the expected utility difference between the left and right options, the temperature (or noise) parameter, λ, and θ which

---

[21] U2-Prelec: the α-parameter is the temperature parameter of the function (generally, if α > 1 utility is S-shaped, if α < 1 utility is inverse S-shaped), and the β-parameter controls the height (or location) of the function's inflexion relative a 45° line across the x- and y-axes of the function.

[22] $U_{2\text{-SCDF}}$: the α-parameter is the power of the function's curvature (if α > 1 utility is S-shaped, if α < 1 utility is inverse S-shaped), and the β-parameter controls the x-axis position at which the function's curvature inverts.

[23] U3-Power: a classic PT interpretation of the power function. Here the α-parameter is the power of the function (generally, if α > 1 utility is S-shaped, if α < 1 utility is inverse S-shaped), the β-parameter accounts for any loss aversion (convexity would be more pronounced than concavity) , and the γ-parameter controls the x-axis position at which the function's curvature inverts.

captured side bias parametrically. The expected utility of each option (EU_{Left}, EU_{Right}), as a function of their probability (p) and the utility function U(m), was given by the functional form:

$$EU(p,m) = p \times U(m) \tag{2-8}$$

The model's best-fitting parameters were estimated by minimizing the following cumulative log-likelihood function:

$$-LL(\theta|\, y) \;=-\left(\textstyle\sum_{i=1}^{n} y_i \times log(P_{Choose\ Gamble}) + \sum_{i=1}^{n} y_i' \times log\left(P_{Choose\ Safe}\right)\right) \tag{2-9}$$

where y and y' indicated a left or right choice (0 or 1), respectively, for each trial i; n was the total trial number for the session.

Again, the best-fitting discrete choice model was selected via BIC comparisons, this time defined on the likelihoods ($BIC_{LL} \;=\; (k \times ln(n)) \;-(2 \times LogLikelihood)$). The parameters estimated in each reward range were also compared using a one-way MANOVA.

2.2.5 *Validating utility predictions from out-of-sample certainty equivalents*

To validate the predictions of the utility functions, CE measures were gathered from binary choices presented outside of the utility estimation sequences – these were then used to corroborate the risk attitudes predicted by the fractile- or DCM-derived utilities. Two of the three monkeys were presented with three sets of four gambles unique to each reward range for which we estimated CEs. We used these 12 CEs to validate the risk-attitude predictions of the utility function estimated in each range. The gambles in the narrow reward ranges had a spread of 0.15 ml, while gambles in the full range had a spread of 0.30 ml – keeping the relative spreads equivalent across the ranges. The gambles' expected values were also, once normalized, centred around the same relative values. In percentage points, each gamble spread over 30% of the reward ranges, and gamble was centred at a value representing 25%, 45%, 65%, or 85% of the reward range (Fig.2-2c).

Taking the difference between the CE of these gambles and their expected value as a proxy for risk attitude ($CE - EV$), the risk-attitude estimated from these CEs were compared with the predictions from the fractile-elicited and discrete-choice utility curves. If the CE – EV metric were positive, it signalled that the animals were risk-seeking. If instead, the measures were negative, the animals could be seen as risk-averse. Because of this, if the utility models imposed an S-shape that was unrealistic (and a consequence of the function used) the $CE - EV$ fits would expose it right away: they would not go from risk-seeking to risk-averse. These measures were repositioned relative the point at which the median (across all testing days) fractile- and DCM-derived utilities predicted that risk-attitudes would reverse (i.e., the point of risk neutrality) and linear regressions were fit to the repositioned $CE - EV$ metrics in order to identify which of the two inflexions proved most reliable in predicting out-of-sample behaviour (fractile or DCM-derived):

$$CE - EV = \beta_0 + \beta_1(EV - inflection) \tag{2- 10}$$

In the model, β_0 indicated the point at which CE measures became risk-neutral, i.e., the inflexion point, and β_1 paralleled the curvature, or steepness of a utility function. The R^2-value associated with both regressions were compared to see which of the two utility estimation procedures most reliably matched out-of-sample behaviour. Put simply, these regressions allowed both the validation of predicted risk-attitudes, and the selection of the better-fitting procedure.

2.2.6 *Defining preference adaptation metrics*

Comparing the utilities estimated from choices in different reward ranges was done in one of two ways: the first, assuming that preferences were fixed and did not adapt to the range of possible rewards in a task; the second, assuming that preferences fully adapted to the reward spread and magnitude of the task at hand (Fig. 2-5). To test for the former, utilities estimated in narrow ranges (i.e., low- and high-range) were compared to the full-range ones. For the latter, the assumption of full adaptation, utilities were compared sequentially - looking for differences in the shape of the utilities between consecutive ranges.

The parametric utility functions all had a unique inflexion point (if at all): a single point where the utility function’s curvature reversed, the maximum of the function’s first derivative. This inflexion identified the precise reward magnitude for which the monkeys’ risk-attitude changed, and, served as a good indicator or where and how the monkeys’ preferences would change depending on the spread and mean rewards in a given range. The inflexion points elicited in different ranges were compared using a Kruskal Wallis test with Bonferroni-Holm corrected post-hoc analysis (Wilcoxon test).

Another metric, the curvature ratio (CR) was defined as the normalized area under the utility functions (the function’s area divided by the total area in each range’s domain). The CR provided a direct, normalized metric of the convexity/concavity interplay of daily utility estimates – reflecting overall risk attitude to a greater degree than inflexion points. A linear utility function (i.e., a diagonal line) would have a CR of 0.5, as would perfectly symmetric S- or inverse S-shaped utilities. A CR above 0.5 indicated that the functions fell above the diagonal and was mostly predicting of risk-averse choices; conversely, a CR under 0.5 reflected more risk-seeking choices. The CRs measured in the different range, were also compared using a Kruskal Wallis test followed by pairwise Wilcoxon rank-sum comparisons (Bonferonni-Holm corrected).

A final series of metrics, defined as adaptation coefficients, allowed for the quantification of relative changes in CRs between utilities that had been estimated in consecutive reward ranges. A sequential adaptation coefficient (SAC) was calculated as:

$$SAC = \frac{\left(\int_{min}^{max} U_n(m)\,dm - \int_{min}^{max} U_{n-1}(m)\,dm\right)}{\int_{min}^{max} U_{n-1}(m)\,dm} \tag{2-11}$$

and it captured changes in the median utility of a given reward range n ($U_n(m)$), where m represented every reward between the minimum and maximum rewards in the tested range, relative to the median utility function in the previous range n-1 ($U_{n-1}(m)$). Since all parametric functions were defined from 0 to 1, comparing the area under each curve gave us a direct measurement of the difference between the utilities that captured preferences in consecutive reward ranges.

A second coefficient, the general adaptation coefficient (GAC), compared the utility of low- and high-reward ranges to the utility estimated from a monkey's full reward range. The GAC placed the narrow-range utilities (i.e., the low and high range ones) relative to the shape of the full-range's utility function. That is, a GAC of 0 would indicate that the narrow-range utilities are but segments of a fixed full-range one, whereas a GAC of 1 suggested that utilities kept a similar form but fully shifted to represent preferences in the new range (Fig 2b). For any GAC where 0 < GAC < 1, utilities had partially adapted. To calculate this, narrow range utilities were rescaled to map onto the full range ones: the maximum value of the low-range became the utility value of the full-range utility at 0.5 ml, and the utility value of the full-range utility at 0.5 ml became the minimum value of the high-range (see Fig. 2-5a). Then, the median utility of the full range (U_{full}) was rescaled (into U_{adapt}) to match the domain and image of narrow range utilities ($U_{\text{low-range}}$ and $U_{\text{high-range}}$). The GAC was given by

$$GAC = \frac{\left(\int_{min}^{max} U_{partial}(m)\,dm - \int_{min}^{max} U_{full}(m)\,dm\right)}{\left(\int_{min}^{max} U_{adapt}(m)\,dm - \int_{min}^{max} U_{full}(m)\,dm\right)} \qquad (2\text{-}12)$$

where min and max are the minimum and maximum reward magnitudes in a narrow range condition. A GAC of 1 signalled full adaptation while a GAC of 0 indicated that no adaptation had taken place. Crucially, the GAC metric took no account of the order in which reward ranges were tested; it relied instead on full-range utility function as a comparison template.

2.3 Results

2.3.1 Experiment design and behavioural performance

Using a binary choice paradigm, I explored the role of contextual reward statistics in shaping monkeys' preferences. On each day of testing, the animals were presented with choice sequences from which their propensity to take risk as a function of the rewards offered was measured. The range of rewards that the animals could expect would be kept constant over consecutive days, then suddenly change to a new distribution of rewards.

With no change to the task other than the reward magnitudes presented and delivered, the monkeys experienced periods of relatively low rewards, periods of relatively high ones, and periods where they saw a mix of both. The aim was to understand how these different ranges informed choice preferences (if at all) and to see how the animals' utility functions shifted with the tested ranges' new reward statistics (Fig. 2-5). To do so, three rhesus macaques were presented with sets of binary choices from which utility functions were either estimated or validated through out-of-sample measurements of risk attitude. In utility estimation sequences, utility measurements were derived from the choices that monkeys made between sets of gambles and safe rewards. Using the fractile method (see *methods*), utilities were derived from the certainty equivalents (CEs) of specific sets of gambles (the safe rewards that were subjectively equivalent to the gamble). In validation sequences, the monkeys' risk preferences were measured directly using the certainty equivalents of out-of-sample equivariant gambles. These measurements were then used to confirm the utilities estimated in elicitation sequences (Fig. 2-5c).

Each experimental range was held constant for a continuous number of days, and on each day the monkeys were presented with either a utility estimation sequences, an equivariant gamble sequences (out-of-sample validation sequences), or both. Monkey A experienced a low range for 22 days (0 ml to 0.5 ml), a full range of rewards for 31 days (0 ml to 1.0 ml), and a high range of rewards for 17 days (0.5 ml to 1.0 ml). Monkey B experiences a low range for 33 days, then 19 days of high range, followed by 18 days of full range. Monkey C, quite uniquely, offered a dataset with a longer timescale. He experienced the full range of 0.1 ml to 1.3 ml of reward for 14 days then switched to a low range of 0 ml to 0.5 ml for 54 weeks. After this, his preferences were measured over 43 days.

For each reward range, sets of daily utilities were estimated using the fractile method. The way reward magnitudes (CEs) mapped onto these utilities (once normalized to the minimum and maximum rewards in a range) could then be compared within and between the different rewards ranges. To do so, and because utilities were defined from 0% to 100% regardless of their range, the CEs were normalized relative the maximum and minimum magnitudes in the appropriate reward range (Fig. 2-3). As expected, higher utility values mapped onto higher reward magnitudes (higher CEs), but the way in which they did so differed markedly depending on the range. The same utility levels (12.5%, 25%, 50%, 75% and 87.5%) in different reward ranges did not map onto the same relative magnitudes (i.e., normalized CEs). I confirmed this statistically, using a two-way ANOVA with the main factors being the utility level tied to individuals CEs and the reward range from which they had originated. The ANOVA confirmed that there was a significant main effect of utility level on the value of the estimated CEs (monkey A: F(4,295) = 64.301, p = 4.812 × 10^{-39}; monkey B, F(4,192) = 50.51, p = 4.107 × 10^{-39}; monkey C: F(4, 295) = 609.547, p = 3.254 × 10^{-141}). The range in which utility-specific CEs had been estimated also had a significant main effect on the value of the estimated CEs (monkey A: F(2,295) = 356.415, p = 1.991× 10^{-79}; monkey B, F(2,192) = 8.994, p = 0.003× 10^{-3}; monkey C: F(1, 295) = 16.204, p = 7.235× 10^{-5}). Together, these corroborated what we could see graphically (Fig. 2-3): higher CEs correlated with higher utilities in all ranges, but these CEs were all relatively lower once a shift from low- to full- or high-range had occurred. Supporting the two other main effects, there was a significant interaction effect of utility level and range on the estimated CEs, in two of the three monkeys (monkey A: F(8,295) = 1.156, p = 0.326; monkey B, F(8,192) = 5.217, p = 1.829× 10^{-5}; monkey C: F(4, 295) = 8.488, p = 1.707× 10^{6}). That is, the steepness of the utility-CE pairings changed between the different reward ranges – utilities in different ranges seemed to follow different functional shapes, rather than a simple shifting and rescaling,

2.3.2 S-shaped utilities best fit monkeys' behaviour

Parametric utility functions were fitted to the daily utility measurements to better compare and understand the relationship between the utilities estimated in each range. To do so, several different functional forms of utility were first compared; the most reliable function was then used for all further analyses. Power functions are commonly used to model utility functions. I therefore fit a 1-parameter power ($U_{1\text{-Power}}$), 2-parameter CDF of a two-sided power ($U_{2\text{-SCDF}}$), and a 3-parameter anchored power functions ($U_{3\text{-Power}}$) to the monkey's CE-utility pairings. In addition to power-type functions, I looked at functions typically reserved for probability distortion modelling (Ferrari-Toniolo et al., 2019; Stott, 2006): the 1-parameter Tversky function ($U_{1\text{-Tversky}}$), and the 2-parameter Prelec ($U_{2\text{-Prelec}}$) – two functions that could readily take on the s-shape prescribed by PT. All functions mapped reward magnitudes onto utility values from 0 to 1 (i.e., 0% to 100% of normalized utilities), and all but the 1-parameter power function could capture risk-seeking and risk-averse behaviour, as well as any inversion in the monkeys' risk attitudes within a reward range.

Because of the fractile method's reliance on aggregate, chained datapoints (Farquhar, 1984; Machina, 1987), utility functions were also fit using a discrete choice model (DCM) applied to individual, rather than aggregate, choices (Eq. 2-7). In line with the fractile-derived utilities, and because previous experiments with the same monkeys had identified negligible probability distortions for $p = 0.5$ (Stauffer et al., 2015), choices in the model were then predicted based on the choices' expected utilities (probabilities were treated as objective). The parameters that best described individual choices in each model were estimated through maximising the cumulative log likelihoods of the DCMs defined on individual experimental sessions (Eq. 2-9; see *methods*).

To select the utility function that best described both the CEs and individual choices Bayesian information criteria (BIC) from all fitted models - the model with the lowest median BIC would thus represent the best fitting model. Of the five tested utility functions, the 2-parameter Prelec proved most reliable in fitting both forms of data (Fig. 2-4a,b). Though the model is normally reserved for probability distortion models, it presented the lowest BIC_{RSS} score as derived from the residuals of fractile-derived utilities (significantly so, Friedman test; monkey A: $F_r(4,240) = 177.154$, $p = 3.046\times 10^{-37}$; monkey B: $F_r(4,168) = 140.780$, $p = 1.903\times 10^{-29}$; monkey C: $F_r(4,220) = 120.800$, $p = 3.604\times 10^{-25}$), and the lowest BIC_{LL} score as derived from the log likelihoods of the discrete choice fits in 2 of 3 monkeys (Friedman test; monkey A: $F_r(4,240) = 219.091$, $p = 2.327\times 10^{-45}$; monkey B: $F_r(4,168) = 186.469$, $p = 2.221\times 10^{-38}$; monkey C: $F_r(4,220) = 180.020$, $p = 5.298\times 10^{-37}$). In the other monkey (monkey A), the BIC_{LL} of the 2-parameter CDF of the two-sided power distribution and the 2-parameter Prelec proved statistically indistinguishable. From these BIC_{RSS} and BIC_{LL} measures, and because the behavioural predictions from each fitting method generally agreed (Fig. 2-4c), I selected the 2-parameter Prelec function for all further analyses.

2.3.3 Monkeys' risk preferences adapt to novel ranges of rewards

Each fitted utility function provided a pair of parameters that could be compared to those elicited in the same or different reward ranges. The curvature of these utility functions served as a direct indicator of the animal's risk attitude for any given magnitude. Convexity reflected risk-seeking behaviour; concavity signalled risk-aversion. From these parametric functions, three predictions could be made: utilities would either (i) fully adapt to the novel reward ranges, (ii) not adapt and remain constant (i.e., different parts of the same curve), or (iii) utilities would partially adapt in a way that did not solely rely on the current reward range. To test for these predictions, further analyses were split into two sets of hypotheses. One set looked at utilities under the assumption that no adaptation had occurred, the other assumed full utility adaptation between each of the reward range conditions. In the case of the no adaptation assumption, the predictions from utilities on identical reward magnitudes in the narrow range and full range were compared (Fig. 2-5a). For the full adaptation assumption, the utilities from sequential reward ranges were normalized and compared, looking at any differences with the previous range's pattern of risk attitude (Fig. 2-5b). If neither assumption proved accurate, then the assumption would be that neither full nor no adaptation had taken place – that is, preferences would have partially adapted.

Starting with fractile-derived utilities, comparing the functional parameters elicited in the different reward ranges provided us with a stringent test regarding the full adaptation assumption. In the 2-parameter Prelec function, the α-parameter represented the temperature of the function, while the β-parameter captured the relative height of the curve. If these were identical across conditions, similar patterns of utility reflected preferences regardless of unique reward magnitudes in the different reward ranges. One-way MANOVA analysis on the log-transformed parameters confirmed that this was not the case: there was a significant effect of reward range on the parameters elicited in each condition, for all monkeys (monkey A: $F(2,59) = 34.913$, Wilks's $\lambda=0.454$, $p = 1.116\times 10^{-10}$; monkey B, $F(2,41) = 13.695$, Wilks's $\lambda=0.594$, $p = 2.946\times 10^{-5}$; monkey C: $F(1, 54) = 9.381$, Wilks's $\lambda=0.739$, $p = 3.252\times 10^{-4}$). Specifically, there was a significant differences between Monkey A and B's β-, or height-, parameters (monkey A: $F(2,59) = 67.301$, $p = 2.447\times 10^{-11}$; monkey B, $F(2,41) = 13.695$, $p = 2.946\times 10^{-05}$; monkey C: $F(2,54) = 1.120$, $p = 0.290$), as well as a significant difference in Monkey C's α-, or temperature-, parameters (monkey A: $F(2,59) = 0.434$, $p = 0.513$; monkey B, $F(2,41) = 2.583$, $p = 0.116$; monkey C: $F(2,54) = 18.858$, $p = 6.236\times 10^{-5}$). The utilities, in terms of parameters, differed depending on the range from which they came (Fig. 2-6b).

To explore how these parametric differences influenced utility patterns in a way that was directly comparable between conditions, I compared the position of each utility function's inflexion points – the

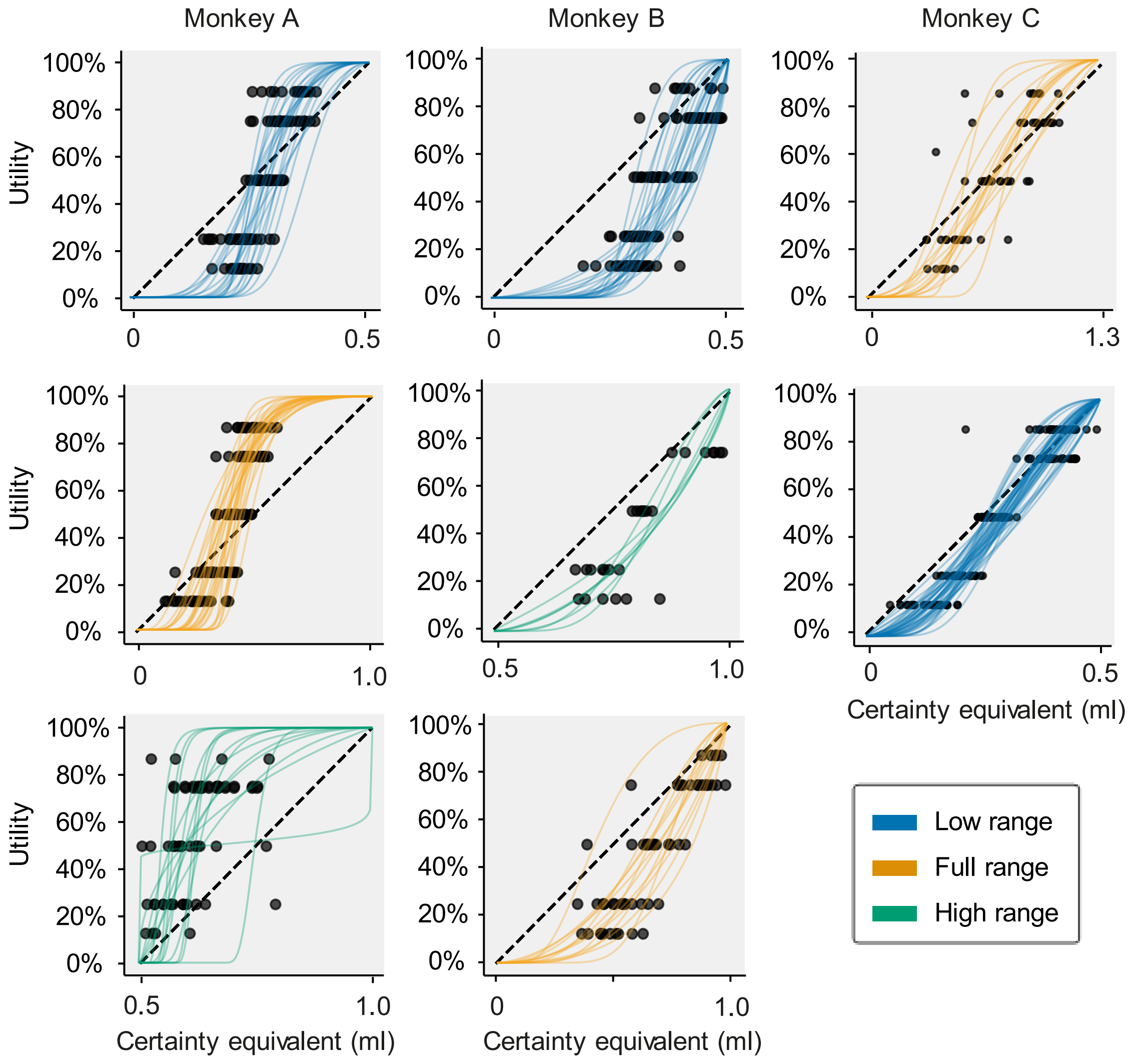


Figure 2-3 | Utility functions elicited from daily fractile procedures. Order of ranges tested is captured vertically. Black dots represent CE-utility pairings elicited in individual experimental sessions using the fractile method; coloured lines are parametric fits ($U_{2-Prelec}$) to daily CE estimates (blue, low narrow range; yellow, full range; green, high narrow-range. Utility fits for Monkey A, from top to bottom, represent 20 days, 26 days, and 15 days. For monkey B, we have 23 days, 7 days, and 13 days. Finally, monkey C has a total of 13 days for the top panel, and 43 days for the lower one. In all cases, convexity of the functional fit signals risk-seeking behaviour, concavity signals risk-aversion.

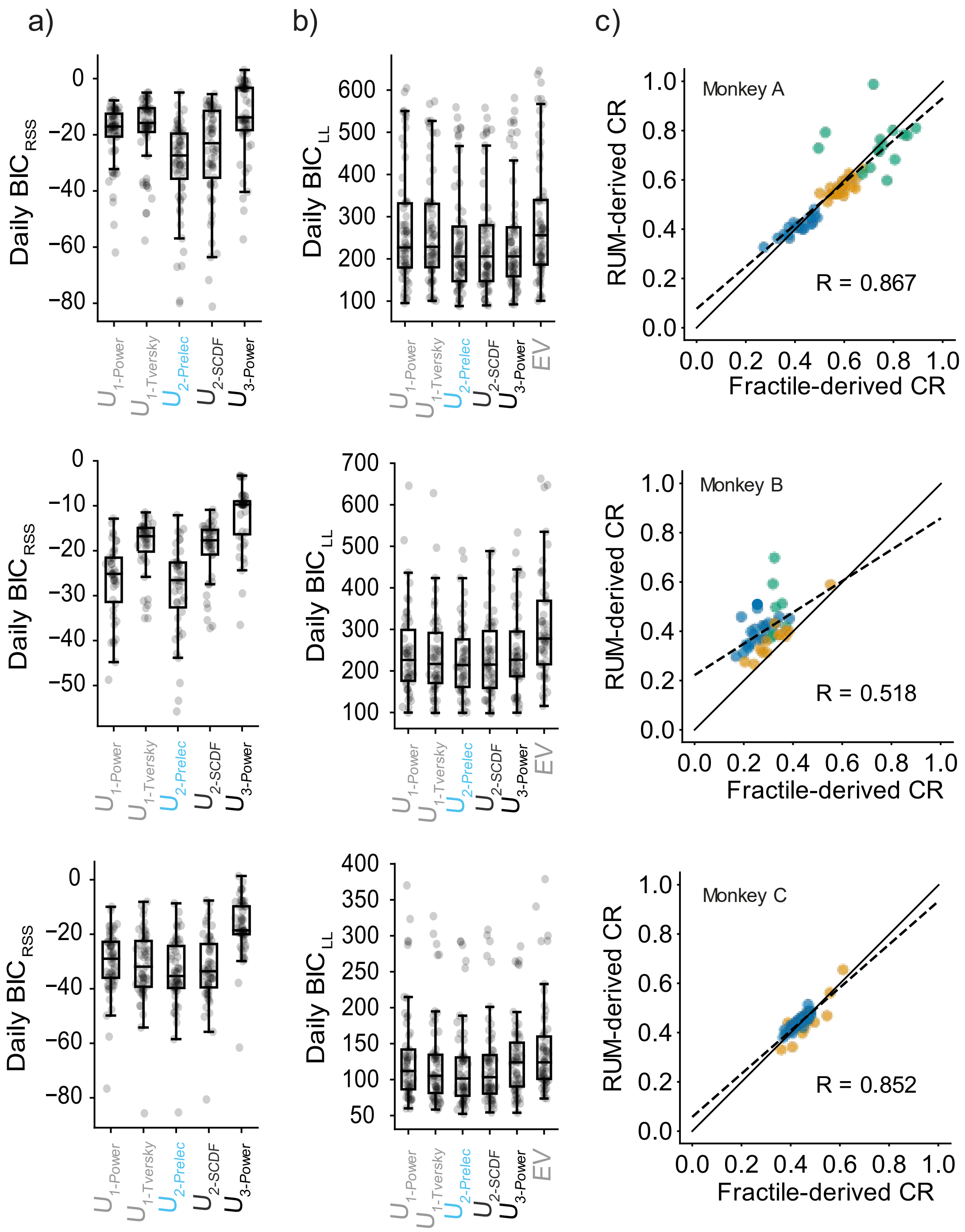


Figure 2-4 | Model comparisons within and across fitting procedures. a) Model selection for fractile-derived utilities. We calculated daily Bayesian information Criterions for each utility function using the orthogonal residuals on each fit (BIC_{RSS}). Lower BIC_{RSS} scores indicated a better fit to the CE-utility pairings, and the 2-parameter Prelec model that was used throughout this study appears in blue ($U_{2\text{-Prelec}}$). b) Model selection for discrete choice utilities. We again calculated daily BIC scores for each utility function, this time using the log-likelihoods estimated to fit each discrete choice models (BIC_{LL}). Lower BIC_{LL} scores indicated better fits between the discrete choice model (DCM) predictions and individual measured choices pairings. Again, the 2-parameter Prelec model that was used throughout this study appears in blue ($U_{2\text{-Prelec}}$), and, in contrast to the fractile-fits, we also compared the various DCMs to predictions based on expected value (seeing if noise alone could explain choices). c) Curvature ratios (CRs) from each fitting procedure correlate. We calculated CRs as the area under the curve of each utility function. Each point represents the CRs from fractile-derived utilities (x-axis) and DCM-derived utilities (y-axis); their colour captures the reward range from which they estimated (blue: low-range, green: high-range; yellow: full-range). Significant positive correlations between the fractile-derived CRs and DCM-derived CRs were found in each of the three monkeys, and we only observed clear differences between the two procedures in monkey B.

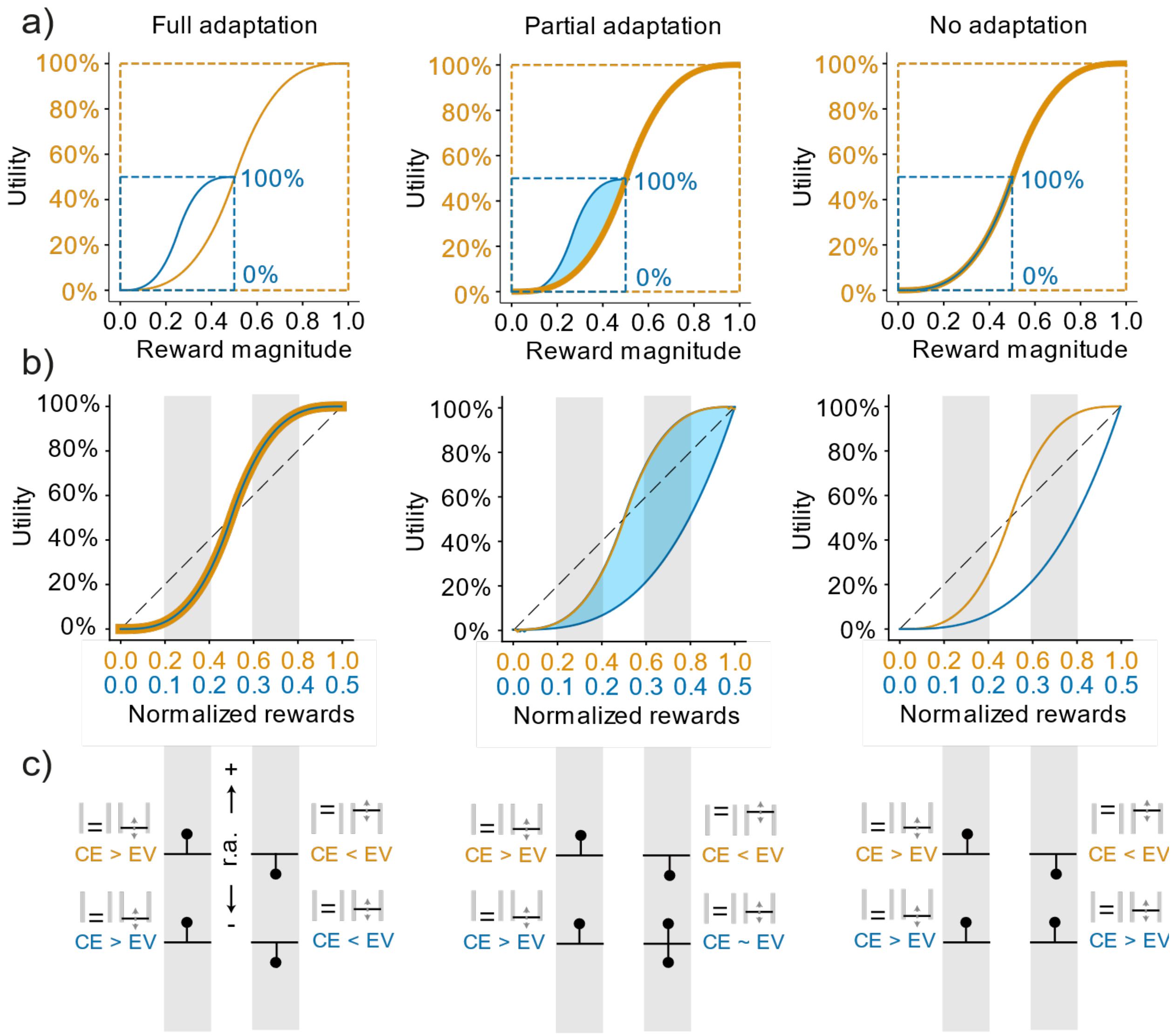


Figure 2-5 | Schematic representation of full-, partial-, and non-adapting utilities estimated in low- and full-ranges of rewards. a) Scaled, identical utility functions in different reward ranges: the utility value of a 0.5 ml reward in the small range (blue curve, 100% utility) is scaled to the utility value of 0.5 ml reward in the large range (yellow curve). From left to right, utilities reshape assuming full-, partial-, and no adaptation. The three possibilities differ mostly in terms of the risk-attitudes exhibited for rewards between 0 ml and 0.5 ml – under full adaptation they should differ, under no adaptation they should not. b) Utilities normalised according to the reward range from which they were estimated. Utilities are set on the same scale by normalizing across the domains of each function. Curves should overlap if utilities adapt fully (left) and fail to do so if there is no adaptation (right). If functions fail to adapt the low range utility is predicted to be identical to the first half of the full range utility curve. c) Predicting the direction of risk attitudes (r.a) from utilities. For an equiprobable gamble made up of the two outcomes that fall at the edges of each grey shaded area, the horizontal black line depicts the expected value (EV) and the black dot above or below signals the direction in which we expect the certainty equivalent (CE). A black dot above the horizontal line signals risk seeking behaviour (or positive r.a.) and a CE of higher value than the EV, and a dot below the line signals risk averse behaviour (negative r.a.). From left to right we again have predictions of r.a. given full-, partial-, or non-adaptive preferences.

reward magnitude at which the behaviour predicted by the utility function flipped from risk-seeking to risk-averse (or risk-averse to risk-seeking depending on the temperature of the utility function). The inflexion crudely summarized choice predictions with a single metric – one that had been previously used to signal monkeys' 'reference-points' (K. M. Chen et al., 2006; Lakshminarayanan et al., 2011). Importantly, since this metric was tied to CE values; one could easily observe if inflexion points fell on similar magnitudes depending on the range in which it had been measured (Fig. 2-5a).

From these inflexion points, the assumption of no adaptation was tested by comparing both within and across-range inflexions. If no adaptation had occurred, the inflexions would be the same within and across the different reward ranges. Testing for the former, i.e., within range differences in inflexion points, no significant pattern of change could be identified – at least for monkeys A and B (linear regression analysis, monkey A: $p_{\text{full-range}} = 0.160$, $p_{\text{high-range}} = 0.472$; monkey B: $p_{\text{full-range}} = 0.270$, $p_{\text{high-range}} = 0.714$; monkey C: $p_{\text{low-range}} = 0.009$). And since monkey C's low range had been tested over a year after changing ranges – the fact that a significant positive slope was identified (the inflexion slowly went up in value over the days of testing) did little to indicate range-swap adaptation. Moving from within range to between range analyses, there were significant differences between the range-specific inflexions for all monkeys (Kruskal Wallis test; monkey A: H(2,58) = 44.281, p = 2.424×10^{-10}; monkey B: H(2,40) = 27.973, p = 8.429×10^{-7}; monkey C: H(1,54) = 28.397, p = 9.881×10^{-8}), which translated into significant pairwise differences (Wilcoxon rank sum) for all but Monkey B's high and full range inflexion points (Fig 2-6a). Simply put, the inflexion points fell on different reward magnitudes for each of the reward ranges. If preferences had truly been non-adaptive, no significant difference across any of the conditions would have been observed.

Since none of the results corroborated the no-adaptation hypothesis, the next step was to test for full adaptation. Rather than comparing the absolute position of the utilities' inflexion points, testing for full adaptation required predicting where inflexion points from a past range would map onto the next range: the assumption being that if the same utility function simply shifted to a new range (i.e., fully adapted), the relative position of the inflexion should be the same. An inflexion at 0.3 ml in the low range, for example, would be placed at 0.15 ml in the full range, and vice versa. However, since an inflexion of 0.3 in the low range would result in a negative magnitude when compared with the high range, inflexions < minimum reward were set at the minimum, and inflexions > maximum reward were set to the maximum. There were significant differences between all consecutive comparisons in Monkeys A and C, and none for monkey B (Fig 2-6a; Wilcoxon rank sum test). From a full adaptation perspective, this suggested that, while Monkeys A and C had not fully shifted their reference to accommodate the new ranges, Monkey B's preferences seemed to follow the same relative pattern across all rewards ranges.

From the inflexion points, the picture that emerged was one of (at least) partial adaptation. That is, the significant differences between the inflexion points corroborated neither the idea of fully- or non-

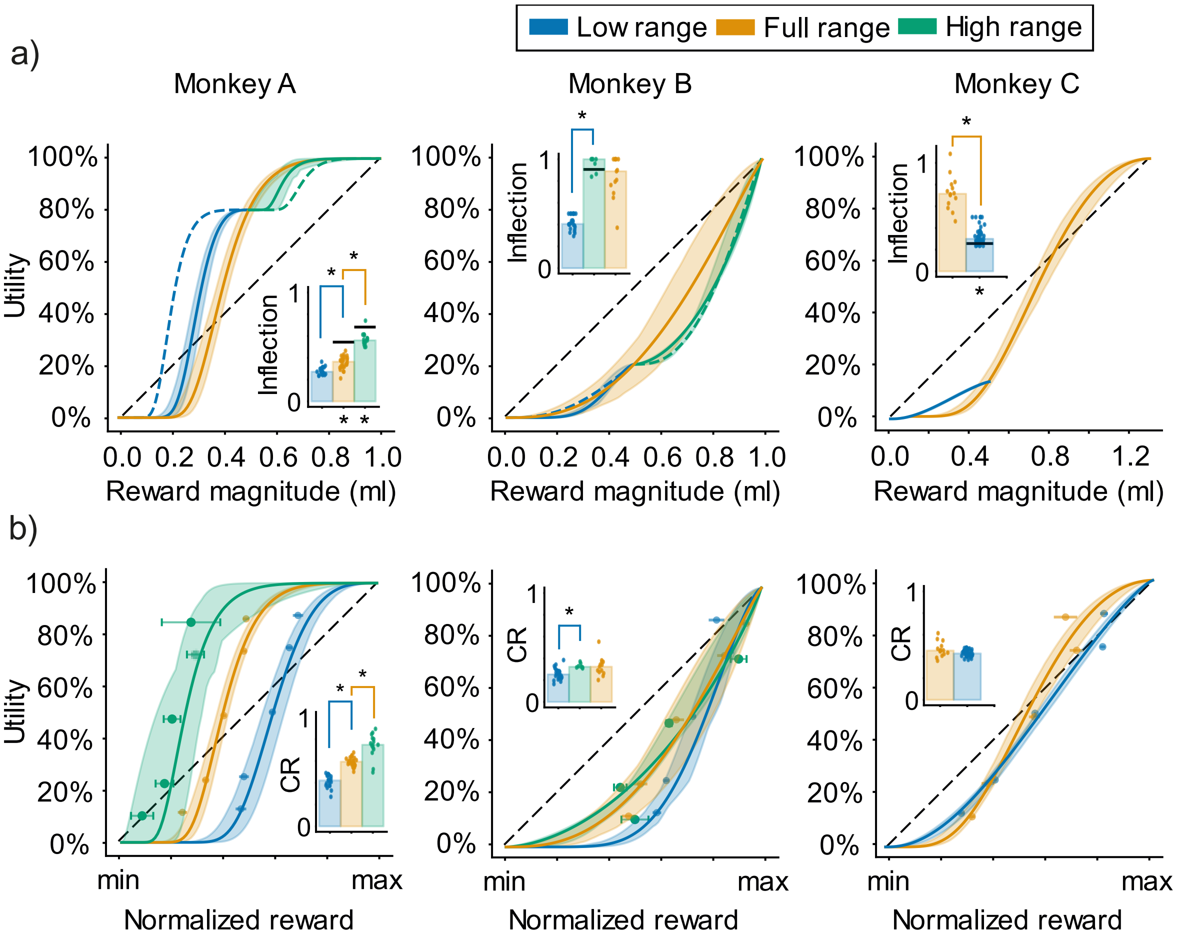


Figure 2-6 | Fractile-derived utilities reflect preferences adaption between the different reward ranges. a) Scaled utilities estimated from fractile-derived CE-utility pairings. Each curve represents the median of daily, range-specific parameter estimates; 95% Confidence intervals were estimated via boostrapping said parameters (random sampling with replacement, n=10000). Dotted blue lines represent predictions full-range utilities predicted to fully-adapt to low-ranges. The dotted green lines represent similar full-adaptation predictions in the high range. Bar graphs represent the median inflection point, i.e., the reward magntiude at which the curve goes from convex to concave – points are daily inflection points. Upper asterisks (*) indicate differences between daily inflection estimates in two sequential ranges (Wilcoxon rank sum test, $p < 0.05$); Lower asterisks (*) indicate significant difference between the median predicted inflection from the previous tested range and the true inflection estimates of the next range (Wilcoxon rank sum, $p < 0.05$). b) Normalized utilities estimated from fractile-derived CE-utility pairings. Each curve is the median of daily, range-specific parameter estimates normalized according to the minimum and maximum rewards in the tested range. Again, 95% confidence intervals were estimated via boostrapping. Points represent mean normalized certainty equivalents ± SEMs for each of the tested range. Bar graphs representmedian curvature ratios (CRs) for each range; the relative concavity of each utility (concave > 0.5; convex < 0.5) – individual points are daily CRs. Upper asterisks (*) indicate significant differences between CRs estimated in sequential ranges (Wilcoxon rank sum, $p < 0.05$). For each panel, blue comes from low-range utilities, yellow from full-range, and green from high-range.

adaptative preferences. Nevertheless, because inflexion points carried no information about the risk attitude that followed or preceded them, the inflexion points could be similar even if the monkeys' choices were not. To counter this, the previous comparisons were repeated using the area under each utility curve – a direct indicator of the convexity/concavity patterns within single utilities. Rather than representing a single point, the area under each curve reflected the order and intensity of risk-seeking or risk-averse behaviour throughout the reward range. Hereafter defined as curvature ratios (CRs, see *methods*), the areas calculated in each range were compared through Kruskal Wallis test (followed by pairwise Wilcoxon rank sum post-hoc tests). The results validated the earlier findings from the inflexion comparisons: sequentially, there were significant differences across ranges for Monkey A and B (monkey A: H(2,58) = 27.973, $p = 8.428\times 10^{-7}$; monkey B: H(2,40) = 12.124, p = 0.002), but there were no statistical differences between monkey C's CRs across conditions (Fig 2-6b; H(1,54) = 1.872, p = 0.171). In essence, while the risk attitudes that monkeys A and B exhibited differed between reward ranges, monkey C seemed to exhibit relatively similar behaviour in the two ranges (albeit with a slightly different inflexion).

To validate these fractile-based comparisons, the full/no-adaptation analyses were repeated using the DCM-derived utilities. Both the inflection points and the CRs of monkey A reliably mimicked earlier findings: significant differences between the ranges meant that inflexion points adapted partially (Fig 2-7a; Kruskal Wallis, H(2,58)= ` 44.504, $p = 2.167 \times 10^{-10}$), but differences in sequential predictions also meant that inflexions were not fully-adaptive (Fig 6a; Wilcoxon rank sum, $Z(45)_{\text{full-range}}$: -5.761, $p = 8.351 \times 10^{-9}$; $Z(40)_{\text{high range}}$: -4.790, $p = 1.661 \times 10^{-6}$). Corroborating the latter, CRs were again found to be significantly different across all range conditions (Fig 2-7b; H(2,58) = -51.342, $p = 7.100\times10^{-12}$). For Monkey B, the DCM-derived inflexion points also behaved like those estimated from fractile utilities: there were significant differences between all but the high and full-ranges (H(2,40) = 31.103, $p = 1.762 \times 10^{-7}$), suggesting that inflexions were not fixed. This was also validated by the finding that there were no significant differences between all consecutive predictions ($Z(29)_{\text{high-range}}$ = 1.103, p =0.270; $Z(20)_{\text{full-range}}$ = 1.941, p = 0.052). In terms of curvature ratios, i.e., test of no adaptation, there again was a difference between the CRs gathered in different reward ranges (H(2,40) = 7.470, p = 0.024), but this time none of the post-hoc pairwise comparisons reached significance once corrected for multiple comparisons (Wilcoxon rank sum; Fig 6b). This meant that monkey B's preferences were much closer to being fully adaptive than not. Finally, Monkey C's results, like monkey A, were consistent across elicitation methods. Inflection points were significantly different between the two ranges tested (H(1,54) = 30.524, $p = 3.297 \times 10^{-8)}$, consecutive inflection predictions were also significantly different ($Z(55)_{\text{low-range}}$ = 2.076, p = 0.03), and CRs were not ($Z(55)_{\text{low-range}}$ = 0.0178, p = 0.897). Inflexions differed, but risk attitudes did not.

Taken together, these results suggest that while no monkey (except perhaps monkey B) demonstrated full adaptation, some form of partial adaptation had occurred across every range in every animal. More specifically, while not fully adapted, monkey A and C's utilities did shift following changes in the task's reward statistics. Their inflexion points moved, but not to the degree predicted by a full shift of the previous range's inflexions. Where the two monkeys differed, however, was in the fact that Monkey C had maintained a very similar CR across conditions – likely due to the time elapsed between the different tests. Monkey B, on the other hand, maintained the relative inflexion predicted across conditions and a similar (though different in fractile-estimates) utility shape.

2.3.4 Predicting range-specific preferences from adapting utilities

While the fractile- and DCM-fits generally agreed on the inflexion of utility functions (Fig. 2-6a, 2-7a), variations in parameter estimates and concavity/convexity patterns (particularly in monkey B; see Table. 2-1) highlighted the need to select the most reliable fitting procedure if quantification of adaptation was the goal.

To address this concern, I compared the risk attitudes predicted by the utilities of each method to real risk attitudes measured in different, out-of-sample choices (i.e., validation sequences). The CEs of equiprobable and equivariant gambles were recorded in each of the reward ranges, and the differences between these CEs and the gambles' expected values (CE – EV) were used to signal the monkeys' risk attitudes. Every gamble had a magnitude spread equivalent to 30% of the respective reward range, and their expected values were anchored at 25%, 45%, 65%, and 85% of the testing range's magnitudes (Fig. 2-2c). If the difference between a gamble's CE and its objective, expected value (*CE - EV*) was positive, it reflected a risk-seeking attitude towards the gamble; if, on the other hand, this value was negative, the animal was said to be risk-averse. These 'validation' measurements were gathered in two of our three monkeys (monkeys A and B).

The *CE - EV* attitude predictions were compared to the risk attitude predictions from the fractile and DCM utility estimates. If the S-shaped pattern of utilities elicited for each monkey were accurate, choices involving magnitudes that fell below the utility's inflexion point should have been risk-prone, while choices above it should have been risk-averse (also validating the S-shape utilities as more than just an effect of the Prelec functional form). I found that this was indeed the case and that CEs in all ranges reflected both risk-seeking and risk-averse behaviour dependent on the relative magnitudes involved (Fig. 2-8a). Then, to identify the best-fitting utility estimation procedure, the CE – EV values were regressed onto the gamble's relative distance from the median inflexions in each range (the distance in EV terms; see Eq. 2-10). In both monkeys, positioning CE – EV values relative the DCM-derived inflexion resulted in a better regression fit than using the fractile-derived inflexions (Fig. 2-8b,c) – the DCM-derived utilities were therefore chosen for further quantification as they represented a more accurate depiction of the monkeys' behaviour.

Table 2-1 | Height and temperature parameters from fractile-derived and DCM-derived utilities

			Utility elicitation method			
			Fractile-derived		DCM-derived	
	Reward range (ml)	N	Median temperature	Median height	Median temperature	Median height
Monkey A	0 - 0.5	20	2.973 95% CIs [2.508, 3.614]	4.798 95% CIs [3.017, 7.387]	2.036 95% CIs [1.820, 2.321]	2.399 95% CIs [2.036, 3.238]
	0 - 1.0	26	4.104 [3.412, 4.572]	0.965 [0.747, 1.256]	3.503 [3.092, 3.588]	1.075 [0.917, 1.287]
	0.5 - 1.0	15	4.606 [2.170, 9.994]	0.155 [0.015, 0.412]	2.961 [2.364, 6.094]	0.167 [0.044, 0.339]
Monkey B	0 - 0.5	23	1.432 [1.312, 1.877]	5.073 [3.975, 6.595]	2.038 [1.875, 2.235]	2.850 [2.689, 3.989]
	0.5 - 1.0	7	0.925 [0.843, 1.423]	1.964 [1.923, 2.410]	1.119 [0.867, 1.592]	1.377 [0.687, 1.807]
	0 - 1.0	13	1.103 [0.943, 1.499]	2.344 [1.878, 3.055]	1.5174 [1.159, 1.988]	2.33 [1.843, 2.647]
Monkey C	0.1 - 1.3	13	1.897 [1.551, 2.043]	1.620 [1.392, 1.751]	1.880 [1.601, 2.060]	1.908 [1.581, 2.161]
	0 - 0.5	43	1.245 [1.166, 1.338]	1.477 [1.429, 1.624]	1.210 [1.191, 1.422]	1.389 [1.343, 1.510]

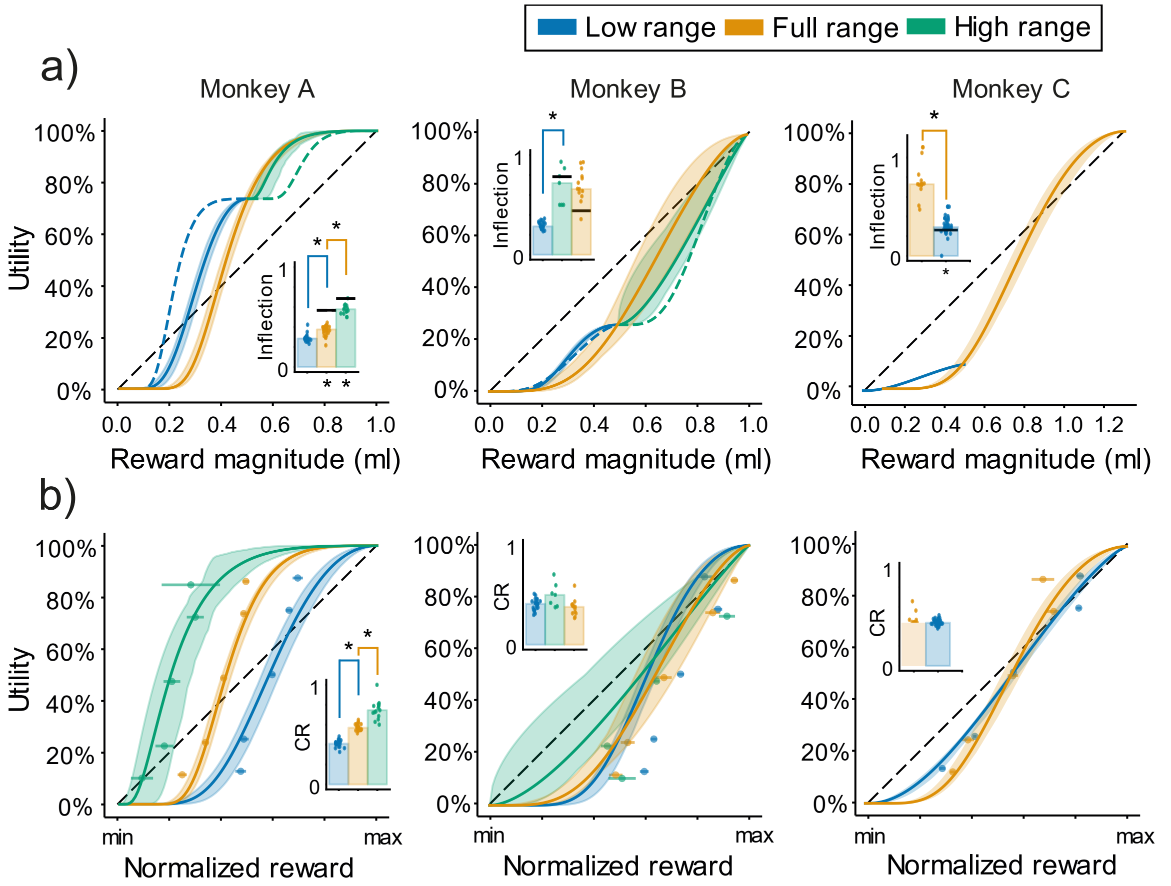


Figure 2-7 | Discrete choice utilities reflect partial preferences adaption between the different reward ranges. a) Scaled utilities estimated from discrete choice models (DCM). Each curve represents the median of daily, range-specific parameter estimates; 95% Confidence intervals were estimated via boostrapping said parameters (random sampling with replacement, n=10000). Dotted blue lines represent predictions full-range utilities predicted to fully-adapt to low-ranges. The dotted green lines represent similar full-adaptation predictions in the high range. Bar graphs represent the median inflection point, i.e., the reward magntiude at which the curve goes from convex to concave – points are daily inflection points. Upper asterisks (*) indicate differences between daily inflection estimates in two sequential ranges (Wilcoxon rank sum test); Lower asterisks (*) indicate significant difference between the median predicted inflection from the previous tested range and the true inflection estimates of the next range (Wilcoxon rank sum). b) Normalized utilities estimated from DCMs. Each curve is the median of daily, range-specific parameter estimates normalized according to the minimum and maximum rewards in the tested range. Again, 95% confidence intervals were estimated via boostrapping (random sampling with replacement, n=10000). Points represent mean normalized certainty equivalents ± SEMs for each of the tested range. Bar graphs represent median curvature ratios (CRs) for each range; the relative concavity of each utility (concave > 0.5; convex < 0.5) – individual points are daily CRs. Upper asterisks (*) indicate significant differences between CRs estimated in sequential ranges (Wilcoxon rank sum). For each panel, blue comes from low-range utilities, yellow from full-range, and green from high-range.

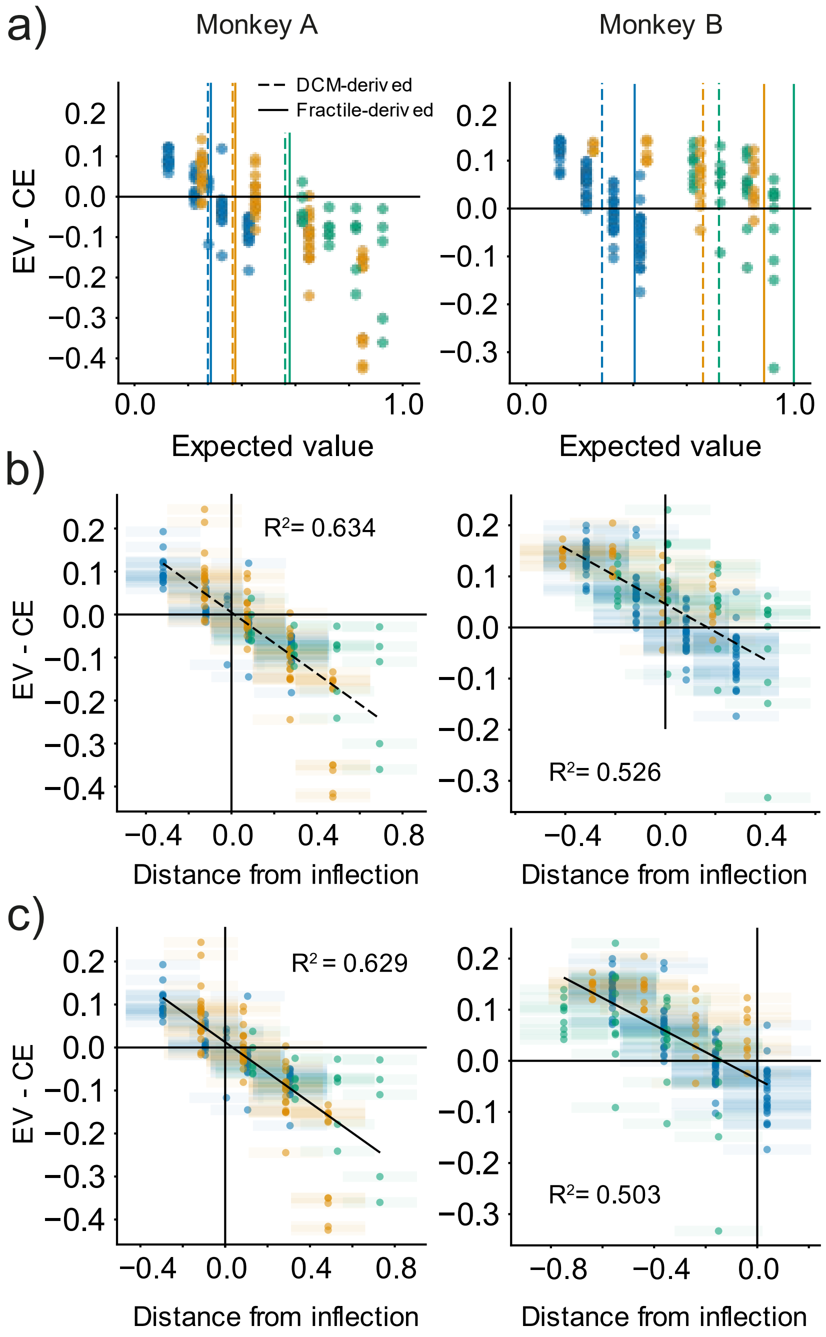


Figure 2-8 | Discrete choice utilities better predict out-of-sample risk attitudes. a) Differences between the certainty equivalent (CEs) and expected value of out-of-sample, equivariant gambles reflects the risk attitudes predicted by utilities. Each point represents a CE – EV measure from individual CE estimates. For CE-EV measures above 0 reflect risk-seeking behaviour, points below 0 reflect risk-averse behaviour. The transition from risk-seeking to risk-averse behaviour should correlate with the inflection points predicted from utility functions: full lines represent the median inflection as predicted from daily fractile-derived utilities; dotted lines represent the median inflection from DCM-derived utilities. b) Discrete choice (DCM) derived inflections (better) predict risk attitudes as measured in out-of-sample gambles. CE – EV metrics positioned as a function of a gamble's EV position relative the median fractile-derived inflection of each range. The x-axis captures the relative difference between the range's inflection point (in ml) and a gamble's EV (in ml). Dotted lines represent linear regression lines across all CE – EV measurements (monkey A: p=1.77 x 10-35; monkey B: p=1.90 x 10-31). c) Fractile-derived inflections predict risk attitudes as measured in out-of-sample gambles. CE – EV metrics positioned as a function of a gamble's EV position relative the median fractile-derived inflection of each range. The x-axis captures the relative difference between the range's inflection point (in ml) and a gamble's EV (in ml). Dotted lines represent linear regression lines across all CE – EV measurements (monkey A: p=5.43 x 10-35; monkey B: p=1.43 x 10-29).

2.3.5 Partial range adaptation shapes monkeys' risk preferences

Two final metrics were defined to quantify the degree to which each monkey's DCM-utilities had adapted between the different reward ranges: a sequential adaptation coefficient (or SAC; Eq. 2-11) and a general adaptation coefficient (or GAC; Eq. 2-12). The former served to quantify how the utilities adapted sequentially as a function of the preceding range of rewards, the latter to position low- and high- range utilities relative to adaptive or absolute full range utilities.

The sequential adaptation coefficients (SAC) represents the percent change in the CRs (the normalized areas under each curve) of successive utilities. It can be used to quantify differences in utilities within a single range, or, in this case, between the median utilities of different ranges. Importantly, the SAC allowed me to quantify utility adaptation on a normalized scale: if utility patterns were fully adapting (i.e., fixed shape regardless of the range), the SAC would gravitate to 0. On the other hand, the SAC would become negative if utilities became more convex (since the area under the utilities would become smaller), and more positive if utilities became more concave. The other coefficient, the general adaptation coefficient (GAC), compared the utility of the low- and high-ranges with the full reward range's utility function (Fig 2b, dashed lines). Using the full-range utility as the 'default' utility shape, the GAC measured how different narrow utilities were – ranging from no or 0% adaptation (i.e., narrow utilities were but segments of an absolute full-range utility) to full or 100% adaptation (the utilities had a fixed form that simply adapted to new ranges). I used DCM-derived utilities to calculate these adaptation coefficients.

Using the SAC to quantify how median utilities changed between ranges, I found that the differences between utilities of monkey A amounted to SACs of 0.37 and 0.35 for the full- and high-ranges, respectively; 0.11 and -0.14 for monkey B's high- and full-range, and 0.04 for monkey C's low range. In utility terms, this meant that Monkey A's utilities predicted behaviour that was 37% and 35% more risk-averse in consecutive ranges. Monkey B also became more risk-averse when going from the low range to the high range but became more risk-seeking again once choosing in the full range. The direction of these changes seemed to reflect the 'position' of the tested ranges relative to the past range the monkeys had experienced. In line with this idea, monkey C had no recent experience with the full-range when low-range utilities were estimated; the measured utilities were thus almost identical.

The GACs calculated for each animal were also very informative in positioning low- and high-range utilities relative to the full range ones (see dotted lines in Fig. 2-7). Monkey A, for example, had a GAC of 0.51 for the small range, and a GAC of 0.21 for the high range. The first essentially meant that the low-range utility was halfway between being but a segment of a fixed full-range utility and being a fully rescaled versions of the full-range utility; the second suggested that high-range utilities were much

closer to being segments of a larger, absolute utility function. For monkey B, low-range utilities matched a GAC of 1.14, i.e., the utilities of the low range had an almost identical shape to those in the

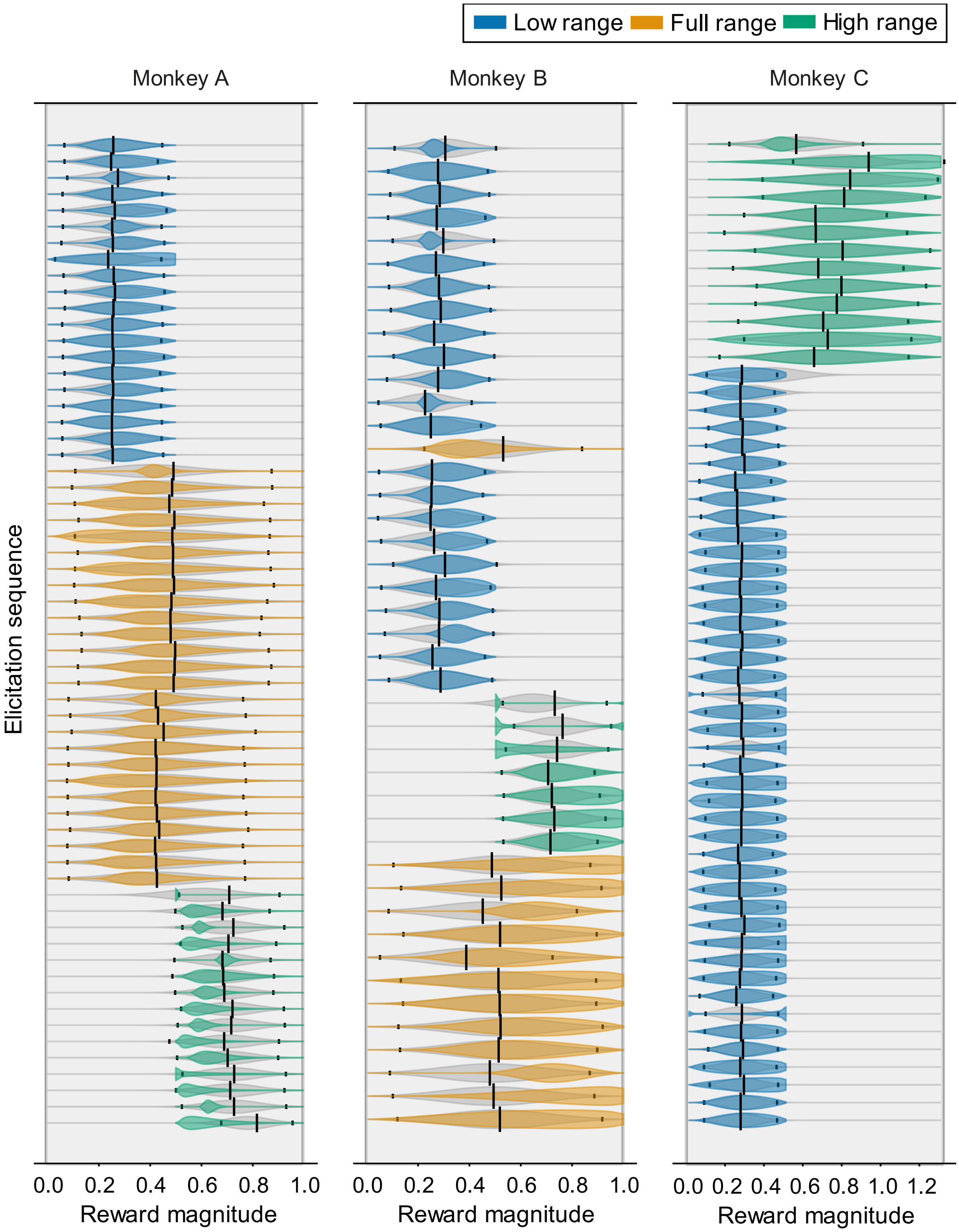


Figure 2-9 | Daily inflections in utilities reflect rewards experienced in past ranges. Each experimental session is represented by a set of horizontal black lines; the first derivatives of fitted utilities appear as the coloured 'violin plots' on the horizontal lines. Black vertical lines indicate the true mean of the rewards experienced by monkeys on individual days – smaller black lines indicate the STD on these means. Grey 'violin plots' reflect the expected distributions of rewards that monkeys 'learned' over past experimental sessions, based on reinforcement learning predictions (Eq. 2-13, 2-14). They are those normal distributions that best fit utilities, as allowed by a Rescorla-Wagner learning rule. Of note, the grey normal distributions are not restricted by 'reward ranges' in the way that utilities are.

full-range, and the high-range utilities had a GAC of 0.69, a bit more than halfway between no- and full- adaptation. Monkey C, corroborating earlier findings, had a GAC between low and full-ranges of 0.87 – they were, for all intents and purposes, identical.

Finally, going back to the original idea that preferences are shaped by one's expectations, I looked at the shape of each DCM-utility relative to the task's daily reward statistics. Though even the initial range's utility inflections never truly followed the task's mean reward (one-sample t-test; monkey A: $t(20)_{\text{low-range}} = 3.849$, $p = 0.001$; monkey B: $t(23)_{\text{full-range}} = 2.534$, $p = 0.019$; monkey C: $t(13)_{\text{high-range}} = 4.267$, $p = 1.103 \times 10^{-4}$), the difference between mean rewards and inflections became markedly larger for monkeys A and B when they were introduced to new reward ranges (Kruskal-Wallis test; monkey A: H(2,58)= ` 40.052, $p = 2.008 \times 10^{-9}$; monkey B: H(2,40)= ` 16.806, $p = 2.242 \times 10^{-4}$). Importantly, the differences were always skewed towards the past range's reward distribution. As reward ranges changed, monkey A and B's references appeared to lag in fully adapting to the new ranges. Monkey C, on the other hand, saw no differences between its two reward ranges ($H(1,54) = 0.021$, $p = 0.884$) – presumably because of the 54-week gap between the two sets of measurements.

To better understand and quantify the lag in fully adapting to current reward, I built a simple reinforcement-learning model that predicted the reward distributions most likely to have shaped monkeys' utilities (Sutton & Barto, 2018). Assuming the 'normal' form and a simple Rescorla–Wagner learning rule, the model then identified the distributions closest to the one captured by monkey's daily utility measures (that is, seeing utilities as the cumulative representation of the reward distribution the monkeys most expected). These distributions' means and standard deviations (STD) were given by the following rule:

$$expected\ mean_i \;=\; expected\ mean_{i-1} + \eta\ (real\ mean_i \;-\; expected\ mean_{i-1}) \qquad (2\text{-}13)$$

$$expected\ STD_i \;=\; expected\ STD_{i-1} + \eta\ (real\ STD_i \;-\; expected\ STD_{i-1}) \qquad (2\text{-}14)$$

where each day's '*expected'* distribution relied on predictions from the previous day ($_{i\text{-}1}$), as well as the learning rate (η) at which monkeys learn from the difference between these predictions ($\textit{expected}_{i\text{-}1}$) and reality ($\textit{real}_{i}$) – the prediction error. Importantly. the first *expected* parameters were assumed to be the statistics that the monkey first observed, because of this as η would get closer to 1, it would indicate that predictions adapted instantly to new ranges; if η was closer to 0, it indicated preferences had relied only on early observations (i.e. the first distribution of rewards that the monkey experience). The functions were fitted by minimizing the sums of square differences between the cumulative distribution function of these curves and the utility of the CEs that had been previously measured using the fractile method.

Though far from a perfect fit, this simple reinforcement model offers some much-needed insight as to the role that expectations played in shaping the monkeys' preferences. Monkeys A, B, and C had learning rates of 0.62, 0.81, and 0.62, respectively; that is, their preferences adapted quickly to new reward distributions, but not fully. The recent past also had a role to play, albeit marginal, in shaping the relative value of rewards. Figure 2-9 illustrates both these ‘expected’ distributions as well as the ‘true’ distributions of the utility curves (as measured by the first derivative of the utility functions). Notice how the expected distributions spill over reward range changes only for the first couple of days. If preferences are built around expectations, then the utilities that best described these preferences point to these monkeys using mostly present but also past information to shape them.

2.4 Discussion

In the present study, I investigated the role that task-specific expectations have in shaping the preferences of macaque monkeys. In line with human research on reference-dependent preferences (Arkes et al., 2008, 2010; Koszegi & Rabin, 2007), monkeys’ risk preferences shifted following changes to the range of rewards they could expect from the task at hand. As the rewards that the task delivered got higher, the reward magnitude at which their risk-attitudes shifted also became higher. Modelling the utility functions that best captured the monkeys’ behaviour, I found that changes in their risk-preferences mimicked the changes predicted in models like Prospect Theory (Kahneman & Tversky, 1979): the points at which utility shifted from convex to concave closely followed what I would consider plausible expectations in the task.

Taking the position of S-shaped utilities as a proxy for the monkey’s expectations, our findings suggest that monkeys partially adapted their preferences to account for new ranges of rewards in a task. While they readily adapted to novel rewards, they did not readily ignore (or forget) reward information that was no longer relevant to the task. Rather than relying solely on the current instalment of the task to build their expectation, the monkeys appeared to also consider the distribution of past rewards – particularly the extremes in a range - in shaping their preferences (i.e., their utility curve). This led to partial, not full, adaptation.

Monkeys A and B, for example, reliably shifted their reference point when possible rewards went from lower to higher magnitudes. When looking at the utility function that best represented their preferences, the monkeys’ utilities appeared to scale instantly to represent the now broader realm of possible rewards. Conversely, when possible rewards were restricted to high magnitudes only (i.e., high-range), the monkeys did not adjust their preferences in a way that accounted for the unavailability of lower magnitudes – even after many days. Where they had previously been flexible in rescaling preferences, the monkeys’ preferences in the high range (where low rewards were never delivered) stubbornly reflected the higher-half of full-range utilities. And while the shift from low to high range seemed to induce partial, almost full adaptation – the shift from full to high range reflected a move along a fixed, absolute utility instead. The data from monkey C, where different reward ranges were tested 54 weeks apart, corroborated this expectation-based interpretation by providing a window on the adaption of utilities after a year. While monkeys A and B experienced every range in the span of just a couple of months, the effects of past high rewards on monkey C would have been minimal. In that respect, it came as no surprise that monkey C's lower range utilities took the form of fully rescaled full-range ones[24].

The idea that preferences adapt to fit the range of possibilities is neither new nor unfounded (Brunswik, 1956; Gigerenzer et al., 1991; Glöckner et al., 2014; E. U. Weber & Johnson, 2008). Indeed, while prospect theory rests on reference-dependence, several newer models mimic RDU in that they claim that the value we imbue our options with rely on the other options we have at our disposal (Hunter & Gershman, 2018; Loomes & Sugden, 2006; Parducci, 2012; Steward et al., 2003; Yaari, 2006). Likewise, it has long been known in psychology and neuroscience that range-adaptation is an inherent feature of the brain (Louie & De Martino, 2013). In sensory systems, for example, neuron's maximize their efficiency by tuning their firing rates to match the distribution of sensory signals (Carandini & Heeger, 2012; Laughlin, 1981) – the same is thought to occur, to varying degrees, in the brain areas that encode value (Burke et al., 2016; Shunsuke Kobayashi et al., 2010; Louie et al., 2015; Padoa-Schioppa, 2009; Tobler et al., 2005; Tremblay & Schultz, 1999). Specifically, and supporting the idea of range-dependent utility, neurons in the primate prefrontal cortex have recently been recorded adapting their firing rate to different reward ranges in a way similar to our monkeys' utility curves. In a study by Conen and Padoa-Schioppa (2019), rhesus macaques only partially rescaled the value of juice rewards relative to the other possibilities in a given block of choices. When recording from neurons in the monkeys' orbitofrontal cortex (an area in the prefrontal cortex often linked to subjective value), the researchers found that the neural code mimicked behavioural measurements in that it partially adapted to match the specific reward distributions of different blocks within the broader context of all past rewards. Crucially, two processes seemed to drive this adaptation: the first, a slow and adaptive learning process about the outcomes one can expect (e.g., reinforcement learning[25], see (Bavard et al., 2018; Rudebeck & Murray, 2014; Wilson et al., 2014); the second, a rapid weighing of rewards relative the decision-makers present context (e.g., normalization[26], see Louie et al., 2013; Hiroshi Yamada, Louie, Tymula, & Glimcher, 2018; Zimmermann et al., 2018)).

Partial adaptation is likely to underlie the brain's ability to maximize 'local' decisions, all while placing these decisions in a much broader context (i.e., relative past experiences; Conen & Padoa-Schioppa, 2019; Fairhall, Lewen, Bialek, & De Ruyter van Steveninck, 2001; Rustichini, Conen, Cai, & Padoa-Schioppa, 2017). When comparing similarly-priced wines, for example, we manage to select our

---

[24] We see a similar effects with Monkey A's utilities that had been previously estimated (years ago) in a different experiment and setup (Genest et al., 2016).

[25] A likely candidate for the adaptive learning process is that of *reinforcement learning*: a process intimately tied with the orbitofrontal cortex and its interactions with the dopaminergic system (for review, see Soltani & Izquierdo, 2019). My own results suggest that reinforcement learning (albeit an over-simplified form) can help explain the role of experience in shaping current preferences.

[26] For the more local, relative encoding of rewards, evidence points to *divisive normalization*: the canonical process by neurons tune their firing rates to match the distribution of available stimuli.

favourite from relatively narrow distributions (similar prices) while still placing our selection relative to a much broader price distribution (our past experiences with wines). It has recently been suggested that this ability to flexibly optimize 'local' decisions while keeping track of past outcomes underlies the formation of cause-and-effect relationships in our thinking (Bavard et al., 2018). If this is the case, then the changes observed in our monkeys' utility functions suggest that the animals truly had an expectation or an internal model about the rewards they could get in the task at hand.

Overall, and in line with the current view from neuroeconomics, I showed that the preferences of macaque monkeys' scale in a way that reflects both inherent properties (and indeed limitations) of the brain and the statistics of the task at hand. Put most poetically by the economists Herbert Simon, our monkeys' decision appeared *"... shaped by scissors whose two blades are the structure of the task environments and the computational capabilities of the actor"* (Simon, 1990, p.7). Perhaps it is time we consider this in building the next generation of predictive economic models.

Chapter 3

Comparing the risky and riskless choice preferences of rhesus monkeys

3.1 Introduction

Whether we are choosing between fruits or vegetables at the supermarket, deciding to jaywalk in the face of incoming traffic, or picking the ideal friends to go traveling with, most of our decisions fall under two categories: some have certain outcomes, some do not. Economists call these risky or riskless decisions (‘risk’ referring to the uncertainty of a choice’s outcome), and - while vastly untested – there is general agreement in economics that peoples’ preferences in one type of situation parallels preferences in the other.

In economics, Expected Utility Theory (EUT) served as the dominant model of risky decision-making until the inception of behavioural economics in the 1970s. Under EUT, a decision-maker’s attitude towards risk was captured by the curvature of their utility function: a mapping of outcomes and rewards onto an internal, subjective metric. Concavity predicted an aversion to risk, while convexity predicted risk-seeking behaviour (von Neumann & Morgenstern, 1944). From this, economists assumed that what accounted for risky preferences in uncertain decisions, i.e., the curvature, could also readily be interpreted as increasing or diminishing preference for larger quantities of a certain outcome (H. F. Barron et al., 1984; Keeney & Raiffa, 1993; Stalmeier & Bezembinder, 1999). Contrasting with this interpretation of EUT, Prospect Theory (PT) highlighted an additional difference between risky and riskless choices through the introduction of subjective probability weightings. Rather than being solely predicted by an individual’s utility curvature, one’s risk-attitude would also vary with their subjective treatment of outcome probabilities (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). In other words, while EUT assumed that risk attitudes derived exclusively from the way in which people value rewards, PT made the case for two components: again, the curvature of utility (i.e., reward valuation), but also the subjective weighting of uncertainty (i.e., distorted probabilities).

PT has since become widespread in the study of risky and riskless decision-making (Camerer et al., 2002; Hertwig & Erev, 2009; Kahneman et al., 1990; Lattimore et al., 1992). But with all the studies on behaviour that make use of PT, there is a remarkable lack of research validating its predictions in both risky and riskless choices; the limitation being that risky utilities (or PT values) are usually measured from choices between risky options (Stott, 2006; Tversky & Kahneman, 1992). This clearly cannot be done in a riskless context, but one interesting avenue has been to compare risky and riskless preferences via introspective metrics. In a study by Stalmeier & Bezembinder (1999), medical patients were asked questions that involved risky outcomes: "*would you rather: live 20 years with a migraine on x days per week (followed by death), or live 20 years with a p% chance of getting migraines y times a week, z times a week otherwise"*; and questions where all options were riskless: "*which difference is larger: the difference between 0 days of migraine and x days of migraine, or the different between x days of migraine and 3 days of migraine"*. Modelling preferences through PT, they found that risky and riskless utilities were identical, and that probability distortion accounted for most of the discrepancy between the risk attitudes predicted by riskless utilities and the risk attitudes measured from risky choices. A similar approach by Abdellaoui, Barrios, & Wakker (2007), this time using money outcomes (gains) rather than medical outcomes (losses), led to the similar conclusion: PT successfully reconciled risky and riskless utilities.

Nonetheless, since the subjects in these studies were generally risk-averse (for gains), it remains to be seen whether PT also reconciles risky and riskless utilities for risk-seeking decision-makers. Additionally, the results of these introspective studies have recently been challenged by a set of studies using a more modern, incentive-compatible approach: the use of time trade-offs as means to study riskless decisions (Cheung, 2016). In these studies, people make choices between larger rewards delivered in the future (with certainty) and smaller rewards delivered now; utilities from intertemporal choices are then compared to those estimated from risky choices. Unlike introspective experiments, however, the majority of the research done on time trade-offs reports discrepancies between riskless, time-discounted utility functions and risky ones (Abdellaoui, Bleichrodt, L'Haridon, & Paraschiv, 2013; Andreoni & Sprenger, 2012; Cheung, 2015; Lopez-Guzman, Konova, Louie, & Glimcher, 2018, but see Andersen et al., 2011) - discrepancies that even probability distortions cannot resolve.

The lack of clear insight as to PT's ability to reconcile risky and riskless choices represents a crucial limitation to the interpretation of the model; particularly as it rapidly becomes the *de facto* model of choice to study animal behaviour and neuroeconomics (X. Chen & Stuphorn, 2018; De Martino et al., 2006; Farashahi et al., 2018; Ferrari-Toniolo et al., 2019; Lakshminarayanan et al., 2011; Marshall & Kirkpatrick, 2015; Stauffer et al., 2015). Simultaneously, since there have been no attempts at reconciling risky and riskless utilities in nonhuman decision-makers, there is no evidence to suggest that either human interpretations can be used. To address this important lacuna, the present study explores the link between the risky and riskless utilities of our close primate relative: the rhesus macaque. I presented monkeys with two types of binary choice trials: risky trials, where monkeys made choices between certain and uncertain juice rewards; and riskless trials, where the monkeys make choices between two certain juice rewards. Like in previous human and macaque experiments, I used the monkeys' propensity to choose uncertain, equiprobable rewards (with varying risk) over several certain ones to estimate the shape of the utility curves that best described their preferences for risk. For riskless choices, utilities were estimated using the random utility maximisation (RUM) framework (for review, see McFadden, 2001). Rather than asking monkeys about the strength of their preferences (as in introspection studies), I measured it directly via the degree of stochasticity they displayed for choices between repeated reward pairs. Since RUM assumes an inherent noisiness on the utilities assigned to reward options, reward pairs where utilities were closer together inevitably led to more stochastic choices. This assumption was used to predict the relative slope of riskless utilities for different reward magnitudes: choices that were more random suggested a relatively flatter curvature of the utility function than choices that were more certain. Importantly, this risky/riskless design addressed two of the most important caveats in human studies: (i) both risky and riskless trials were now incentive compatible (relying on revealed preferences rather than introspection), and (ii) choices were presented in the exact same way for both risky and riskless sequence (no time trade-off).

By parametrically separating the contributions that utility and probability distortion had on the monkeys' risky choices, I found that, just like the human studies had previously shown, risky utilities were closer to riskless utilities once probability distortions had been accounted for. I did not, however, find that these utilities were identical. In fact, after re-estimating risky utilities from a different set of choices (this time with gambles that were both equiprobable and equivariant) I found that more than just a risky/riskless dichotomy, the simple act of holding risk as a constant versus having risk vary led to the monkeys exhibiting significantly different patterns of choice.

Overall, while the hypothesis guiding this study was that PT would successfully reconcile risky and riskless choices, I found that the sole addition of probability distortions to the utility framework did not bring the risky and riskless utilities of macaque monkeys inline. In fact, more than a simple difference in risky/riskless choices, I found that even in risky scenarios, the utilities measured from one type of sequence (varying risk) could not, under PT, predict choices in another type of sequence (constant risk).

3.2 Methods

The premise of this study was to compare the utilities estimated from monkeys' choices in risky or riskless settings. To do so, monkeys were presented with sets of choices that could then be translated into utility metrics. Those utilities measured from riskless choices were compared, first with utilities derived from risky choices, assuming no subjective weighing of probabilities (EUT utilities), then with risky utilities where the contribution of probability distortion had been accounted for separately (PT utilities). Afterwards, since the utilities in each pairing had been found to differ significantly, I added an additional out-of-sample test where I compared the behaviour predicted from these utilities to behaviour in a different set of risky choices. Each of these steps are discussed in further detail in the following sections.

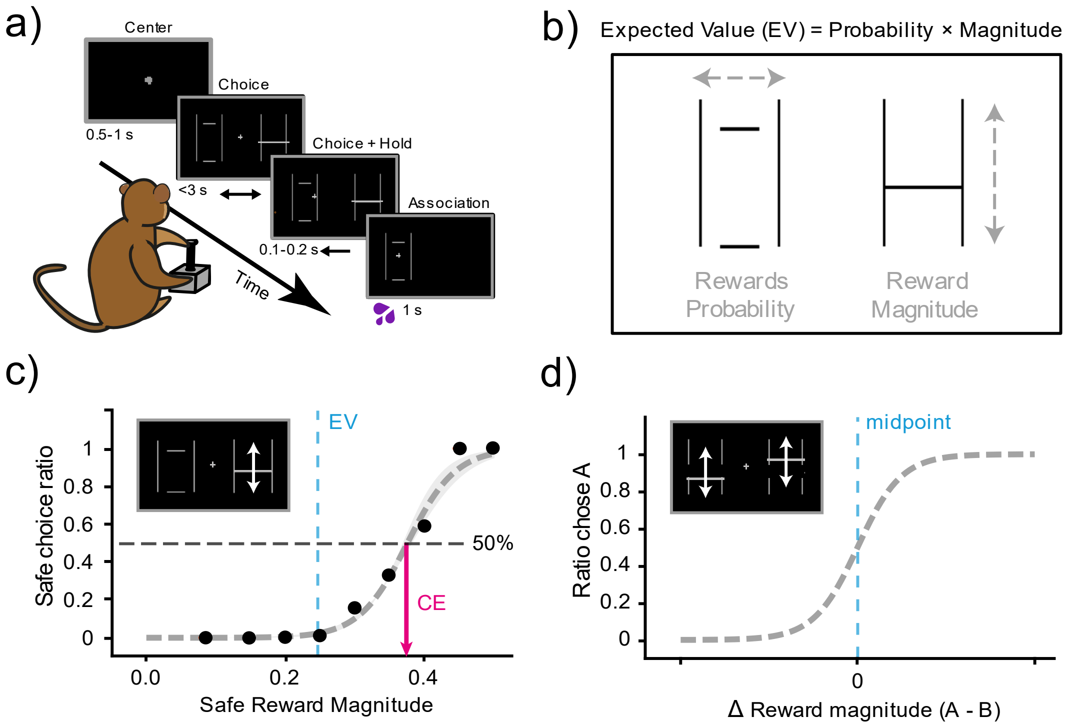


Figure 3-1 | Experimental design and measures of risky and riskless choices. a) Binary choice task. The monkeys chose one of two gambles with a left-right motion joystick. They received the blackcurrant juice reward associated with the chosen stimuli after each trial. Time, in seconds, indicate the duration of each of the task's main events. b) Schema of visual stimuli. Rewards were visually represented by horizontal lines (one or two) set between two vertical ones. The vertical position of these lines signalled the magnitude of said rewards. The width of these lines, the probability that these rewards would be realized). c) Estimating certainty equivalents from risky choices. Monkeys chose between a safe reward and a risky gamble on each trial. The safe rewards alternated pseudorandomly on every trial – they could be of any magnitude between 0 ml and 0.5 ml in 0.05 ml increments. Each point is a measure of choice ratio: the monkey's probability of choosing the gamble option over various safe rewards. Psychometric softmax functions (Eq. 3-1) were fit to these choice ratios, then used to measure the certainty equivalents (CEs) of individual gambles (the safe magnitude for which the probability of either choice was 0.5; black arrow). The solid vertical line indicates the expected value (EV) of the gamble represented in the box. d) Estimating the strength of preferences from riskless choices. Riskless safe rewards were presented against one another, the probability of choosing the higher magnitude option (A) is plotted on the y-axis as a function of the difference in magnitude between the two options presented (Δ magnitude). The differences in magnitude tested were 0.02 ml, 0.04 ml, 0.06 ml, and a psychometric curve, anchored with its inflection anchored at a Δ magnitude of 0, were fit on the choice ratios measured (Eq. 3-2). These functions were fit to different magnitude levels, and the temperature of each curve reflected the strength of monkeys' preferences at each of these different levels.

3.2.1 Task Design and Setup

Two male rhesus macaques (Macaca mulatta) repeatedly chose between two reward options (reward-predicting stimuli) presented on an upright computer monitor. While sitting in a primate chair (Crist instruments), they used a left-right joystick (Biotronix Workshop, the University of Cambridge) to indicate their choice on each trial and received the reward they selected at the end of each of these binary choice trials (Fig. 3-1a).

Reward options took the form of various combinations of reward magnitude and probability, and were represented on the monitor through horizontal lines that scaled, and moved, relative to two vertical 'framing' lines (fig 3-1b). Reward magnitudes were represented by the vertical position of the horizontal lines: 0 ml at the bottom of the vertical frame (1.5ml at the top, and $0 < m < 1.5$ in-between), whilst the probability of receiving said reward was represented by the width of the horizontal lines within the frame. A single, horizontal line that touched the frames at both ends signalled a 100% probability of getting that reward (certain outcome); multiple lines that failed to touch the frames indicated probabilistic gambles of probability $0\% < p < 100\%$ (Fig. 1a). The monkeys were trained to associate these two-dimensional visual stimuli with blackcurrant juice rewards over the course of two years, and both monkeys had previous experience with the task and stimuli before this study. They had both experienced reward probabilities that ranged from 0 to 100% (Ferrari-Toniolo et al., 2019), and reward magnitudes that ranged from 0ml to 1.3ml of juice (see chapter 2). For this study, reward magnitudes were held between 0ml and 0.5ml of blackcurrant juice, and gamble options all had a probability of 50%.

Each binary choice trial began with a white cross at the centre of a black screen, if the monkey was holding the joystick, a cursor would also appear on the screen (Fig. 3-1a). Using the joystick, the monkeys initiated each trial by moving the cursor to the centre cross and holding it there for 0.5-1s. Following this holding period, two reward options appeared to the left and to the right of the central cross (see Fig. 1a). The animal had 3s to convey his decision by moving the joystick to the selected side and holding his choice for 0.1-0.2s - the unselected option would then disappear. The selected option lingered on the screen for 1 s after reward delivery – followed by a variable inter-trial period of 1–2 s before the next trial. Errors were defined as unsuccessful central holds, side selection holds, or trials where no choices were made. Each of these resulted in a 6 s timeout for the animal, after which the trial would be repeated (ensuring the elicitation of preferences for each tested option pair). Additionally, all reward options were repeated on both the left and right sides of the computer screen, alternating pseudorandomly to control for any side preference. Both the joystick position and task event times were sampled and stored at 1 kHz on a Windows 7 computer running custom MATLAB software (The MathWorks, 2015a; Psychtoolbox version 3.0.11). I collected on average 423 ± 91 (STD) trials per session over 22 sessions for monkey A, and 338 ± 41 trials over 7 sessions for monkey B. Only trials where the option pair had been repeated at least 4 time were analysed in this study. Data processing and statistical analyses were run in python (Python 3.7.3, SciPy 1.2.1, see Oliphant, 2007).

3.2.2 Revealing preferences for risky and riskless choice

The monkeys' daily reward preferences were measured in risky and riskless choice sequences under the framework of utility maximisation. In risky choice sequences, trials always pit a risky gamble against a safe option –the utility of different reward magnitudes was estimated via the ratio of choices between different gamble and safe rewards. All of the gambles comprised two equally likely reward outcomes (though one could be 0ml). In riskless choice sequences, monkeys were presented with pairs of 'safe' options with a single fixed outcome – I used the ratio of choice between pairs of rewards that always differed by the same magnitude amount to estimate utility.

For risky sequences, utilities were estimated using the fractile-bisection procedure – a method that involves dividing the range of possible utilities into progressively smaller halves (or fractals) and estimating the reward magnitude associated with each of these utility fractals. Simply put, the procedure defined set utility metrics (in this case ½, ¼ and ¾, and 1/8 and 7/8 of the maximum utility, see fig 2ab) for which the corresponding safe rewards were derived (Fig. 3-2a).

Utility values of 0 and 1 were arbitrarily assigned to 0ml and 0.5ml of juice, respectively. Since monkeys only experienced trials set between these reward magnitudes, this constrained all utility estimates between a 0 and 1. Then, in accordance with EUT[27], a utility of 0.5 was assigned to the equiprobable gamble formed of these two magnitudes (0.5 = [0.5 * 0ml] + [0.5 * 0.5ml]). The first step of the procedure involved presenting the monkeys with choices between this gamble and varying safe rewards (in 0.05ml increments), from these, the safe reward that was equivalent to the gamble in utility terms was identified (i.e., the safe reward chosen in equal proportion to the gamble; see Fig 3-1c).

To estimate this safe reward, the following logistic sigmoid curve was fitted to the proportion of safe choices for each of the gamble/safe pairing:

$$P(ChooseSafe) \ = 1/(1 + \ e^{-\left(\frac{SafeReward_{ml} \, - \, x_0}{\sigma}\right)} \qquad (3\text{-}1)$$

Where probability that the monkeys would choose a safe reward over the 0.5 utility gamble ($P_{(ChooseSafe)}$) was contingent on the safe option's magnitude ($SafeReward_{ml}$) and two free parameters: x_0, the x-axis position of the curve's inflection point, and σ, the function's temperature. Importantly, this function's inflection point represented the exact safe magnitude for which the monkeys should be indifferent between the set gamble and a given safe reward. The x_0-parameter could thus be used as a direct estimate of the gamble's certainty equivalent (CE), or, put simply, the safe reward equivalent to a utility of 0.5.

---

[27] The fractile procedures does not separate the effects of probability distortion from the effects of utility on choices. It therefore relies on EUT's formulation that EU = p x *u*(ml). This limitation will later be addressed using a PT-derived discrete choice model with which the individual contributions of probability distortion and utility can be separated.

Only sequences that contained a minimum of three different choice pairs (repeated at least 4 times) were used in the elicitation of CEs.

From the CE identified as the 0.5 utility value, two new equiprobable gambles were created representing utility values of 0.25 (¼ of the utility range) and 0.75 (1/4 and ¾ of the utility range, respectively). Of the two new gambles, one was set between 0ml and the first CE's ml value, the other was set between the first CE and 0.5ml (Fig. 3-2b). The CE elicitation procedure (logistic fitting, Fig 3-1c) was repeated for each of these gambles. Crucially, gamble/safe pairings for both gambles were interwoven in the same sequence – to ensure a similar spread in the presented rewards.

After eliciting the CEs of these gambles, the estimation procedure was repeated one final time with the new CEs as the upper or lower gamble outcomes. Here, the fractile procedure would automatically terminate if no safe rewards could fit between the outcomes of the new gambles; this would occur if the animal was particularly risk-seeking or risk-averse. If this was the case, utilities of 0.25, 0.5, and 0.75 would be mapped onto the appropriate reward magnitudes and the elicitation sequence would end. If, instead, the three fractile steps were successful, the procedure would result in a mapping of five utilities, 0.125, 0.25, 0.5, 0.75, and 0.875, onto five safe rewards. Only sequences where at least 3 utility points were successfully identified were used in the study (monkey A, 22 sessions; monkey B, 7 sessions).

For riskless choice sequence, choice ratios between pairs of safe options were measured - this time looking at the likelihood of a monkey choosing the high magnitude option over the lower magnitude one (Fig. 3-1d). The range of juice rewards (0.05 ml to 0.5 ml) was divided into sets of 0.05 ml increments and safe-safe pairs centred on these magnitude increments. For each increment, I defined three sets of safe-safe choices where each pairing differed by 0.02ml, 0.04ml, or 0.06 ml. The small size of these differences ensured that choices would be stochastic. These differences are hereafter defined as 'gaps', i.e., safe-safe pairings of fixed differences, where three sets of gaps were anchored at each incremental 'midpoint'.

The likelihood of choosing the higher magnitude option in different gap-midpoint pairings was used to infer the shape of the monkeys' utility functions (Fig. 3-3a,b,c). Specifically, the difference between the likelihoods of choosing the better options, at different midpoints, reflected the separability of the utility of different reward magnitudes. Under RUM, the degree of certainty with which choices are made (i.e., the closer choice ratios are to 100%) directly correlates with the separability of the noisy utilities that correspond to each option in a choice (see this thesis' **Methods Interlogue**). Practically speaking, this implies that, looking at repeated choices between two set magnitudes, a decision-maker with a flatter utility function should exhibit more stochasticity in their choices (i.e., less precision) than a decision-maker with a steeper utility (i.e., more precision). Changes in choice ratios between sequential midpoints, as averaged across gaps, could therefore be used as a proxy for a monkeys' utility slope.

To estimate these RUM-compliant utilities, logistic curves were fitted to the likelihood of choosing the better option (for the three gaps) at every midpoint level (Fig 3-3a):

$$P(ChooseHigher) = 1/(1 + e^{-\left(\frac{Gap_{ml}}{\sigma}\right)}) \tag{3-2}$$

Unlike for CE estimation, this logistic function captured the likelihood of choosing the high-magnitude option (in a safe-safe pairing) contingent on the gap between the two options (Gap_{ml}) and σ, the logistic function's temperature. Just as is the case for CE estimation however, the utility estimates relied on aggregate choices between multiple reward pairs. The logistic fit also highlighted sequences where monkeys would not follow even the most basic principle of rational choice: weak stochastic dominance (picking an objectively lower outcome). Choices where this was the case were removed from all future analyses: that is, when the estimated temperature parameters of logistic fits were negative (i.e., the larger the gap, the lower the likelihood of choosing the better option).

Where logistic fittings were successful, the functions were used to estimate the higher-lower choice ratio, at each midpoint, assuming a magnitude gap of 0.03ml (Fig. 3-3a). Then, the inverse cumulative of a logistic probability density function (centred at 0 with variance = 1) was used to estimate the distance, in utility terms, between the two magnitudes in the 0.003ml gap (Fig. 3-3b). In other words, these 0.03ml gaps were placed onto a shared scale (i.e., random utilities) through the assumption that, on each trial, the probability that the monkeys would pick the better reward (x_i) was given by:

$$P(x_i) = P[U(x_i) \geq U(x_j)], for\ all\ i \neq j \tag{3-3}$$

$$P(x_i) = P[u(x_i) + \varepsilon_i \geq u(x_j) + \varepsilon_j]. for\ all\ i \neq j \tag{3-4}$$

$$P(x_i) = P[u(x_i) - u(x_j) \geq \varepsilon_j - \varepsilon_i], for\ all\ i \neq j \tag{3-5}$$

In this form, the probability of choosing x_i rather than x_j was given by the probability that the difference in the true utilities of x_i and x_j was greater or equal to the noise on x_j (ε_j) minus the noise on x_i (ε_i). From this, it followed that the distribution of noise differences could be used as a predictor of the distance between the two true utilities ($u(x_i)$ and $u(x_j)$). Because of the assumption of constant noise, the probability of choosing x_i over x_j would be directly proportional to the distance between the true utility of two options. In accordance with McFadden's formulation (McFadden, 1974a, 2005; Stott, 2006), I assumed that the distribution of error differences ($\varepsilon_j - \varepsilon_i$) took a logistic form:

$$P(x_i) = \frac{1}{(1+e^{-\Delta utility})} \tag{3-6}$$

and then used the inverse of this logistic distribution's CDF to estimate the difference in utilities ($\Delta utility$) between the hypothetical 0.03ml reward gaps (fig 3-3c) - essentially the slope of the monkeys' utility function at every midpoint. The cumulative sum of these slopes provided an estimate of the utility at each midpoint.

3.2.3 Modelling risky and riskless choices with Prospect Theory

Because the utilities measured from aggregate behaviour did not dissociate between the effects of probability distortion and utility on choices (i.e., they were EUT utilities rather than PT ones), parametric utility functions were re-estimated from individual choices using a discrete choice model that could account for the effects of both, separately. This placed both utility metrics on a common and comparable scale, and, most importantly, it allowed for the inclusion of probability distortion as an additional contributor to the monkeys' preferences.

As in most discrete choice models (and in line with the aggregate RUM metric), a logit function (softmax) was used to represent noise in the decision-making process. The probability of the monkey making either a left or right option was therefore given by:

$$P_{chooseLeft} = \frac{1}{(1+e^{-\lambda(V_{Left}-V_{Rght}-\theta)}} \qquad (3\text{-}7)$$

Where the probability of choosing the left option is a function of the difference in value between the left and right options, the temperature (or noise) parameter, λ, and the side bias parameter θ. The value of each option (V_{Left}, V_{Right}) took on the functional form prescribed by PT:

$$V(p,m) = w(p) * u(m) \qquad (3\text{-}8)$$

where the value of an option was the product of a power-function distorted probability:

$$w(p) = p^{\rho}, \quad (\rho = 1\ if\ choices\ are\ riskless) \qquad (3\text{-}9)$$

and the utility of the option's outcome was prescribed the CDF of a two-sided power distribution (Kotz & Dorp, 2010):

$$u(m) = \begin{cases} \kappa \times \left(\frac{m}{\kappa}\right)^{1/\alpha}, for\ 0 \le m \le \kappa \\ 1 - (1-\kappa) \times \left(\frac{1-m}{1-\kappa}\right)^{1/\alpha}, for\ \kappa < m \le 1 \end{cases} \qquad (3\text{-}10)$$

In the case of the probability distortion function, the ρ-parameter prescribed either an overweighing (ρ >1) or underweighting (ρ<1) of an outcome's probability. The utility prescribed by the 2-parameter CDF of a two-sided power distribution assigned utility measures as a function of an α-parameter (the function's temperature) and an inflection point κ, where the curvature of the utility function would invert. Importantly, for the utility function, the κ was bounded by the range of outcome magnitudes experienced by the monkeys (0ml and 0.5ml) – and each outcome magnitude was normalized onto a 0-1 scale.

Each of these parameters was fit to single-choice data by minimizing the sum of negative log-likelihoods defined on the model as:

$$LL(\theta|\,y) = -\left(\sum_{i=1}^{n} y_i * log\left(P_{Choose\ Left}\right) + \sum_{i=1}^{n} y_i' * log\left(1 - P_{Choose\ Left}\right)\right) \qquad (3\text{-}11)$$

For each individual choice trial, i, y and y' indicated a left or right choice (1 if yes, 0 if no), n was the total trial number for the session, and $P_{Choose\ Left}$ was the output of the earlier logistic function (Eq. 3-3). This discrete choice analysis was restricted to choice sequences previously deemed appropriate for the aggregate preference estimations described in earlier sections.

3.2.4 Statistical comparison of risky and riskless choices

Estimating utilities through discrete choice modelling allowed for the comparison of the functional parameters that best described the monkeys' decisions in risky and riskless conditions, and to explore the unique contributions of both magnitudes (utility) and probabilities (probability distortion) in a way that aggregate, non-parametric measures did not permit – mostly because of scaling.

Because the logit function's λ-, and the utility's α-parameters were asymmetrically distributed (with values <1 accounting for as much change as values >1), these were log-transformed before proceeding with any comparison. Then, the parameters elicited in risky choice sequences were compared to those estimated from riskless sequences using a one-way multivariate analysis of variance (or MANOVA) whereby the main comparison factor in the analysis was the risk-riskless condition described by each set of parameters. Since the probability distortion parameter for riskless choices was constant and fixed at 1, I restricted the MANOVA analysis to the softmax and utility parameters. I then ran additional correlation analyses (Pearson's R) between risky and riskless utility parameters to determine if the parameters in one set of choices could predict those of another.

All parameters were compared independently for each monkey, results were never pooled across animals, and the statistics for each monkey are reported separately. All statistical analyses were considered significant at $p < 0.05$.

3.2.5 Predicting choices from risky and riskless utilities

To better visualize and understand how the different parameters would translate to behaviour, the daily utility and (when applicable) probability distortion functions estimated from risky and riskless choices were used to predict the certainty equivalents (CE) of four different equivariant gambles (Fig 3-6a). Since each daily set of parameters led to four different CEs, these estimates formed distributions that could easily be contrasted with each other and with true, measured behaviour.

Choice sequences to measure these four gambles' CEs were presented to the monkeys on different days – these CEs were then compared to those predicted by the risky/riskless functional estimates. Including the measured CEs, four sets of four CEs (one real, three predicted) were compared using a two-way

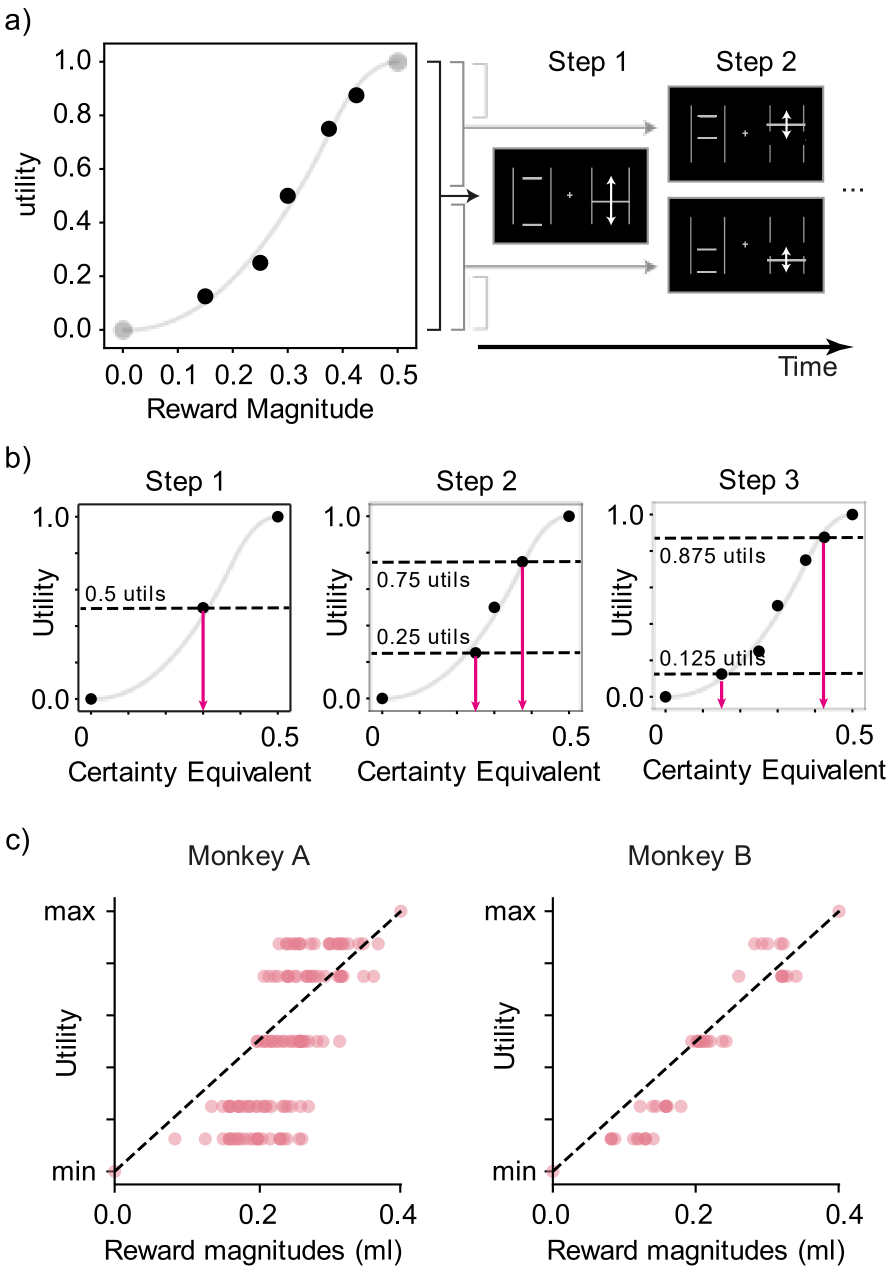


Figure 3-2 | Estimating risky utilities using the fractile procedure. a) Fixed utilities are mapped onto different reward magnitudes. The gambles that monkeys experienced are defined from bisections of the range of possible reward magnitudes. For each step the gambles were held fixed; safe magnitudes varied by 0.05ml increments. b) Estimation of utility using the stepwise, fractile method. In step 1, the monkeys were presented with an equivariant gamble comprised of the maximum and minimum magnitudes in the tested reward range. The CE of the gamble was estimated and assigned a utility of 50%. In step 2, two new equivariant gambles were defined from the CE elicited in step 1. The CEs of these gambles were elicited and assigned a utility of 25% and 75%. Two more gambles are defined in step 3, from the CEs elicited in step 2. Their CEs were then assigned a utility of 12.5% and 87.5%. Parametric utility functions, anchored at 0 and 1, were fitted on these utility estimates (see *methods*). c) Datapoints represent CEs tied to specific utility levels, as estimated through the fractile procedure. Both monkeys exhibit risk-seeking behaviour for low-magnitude rewards, and risk-averse behaviour for high-magnitude ones. The data represents individual utility estimates gathered over 22 sessions for monkey A, and 7 sessions for monkey B.

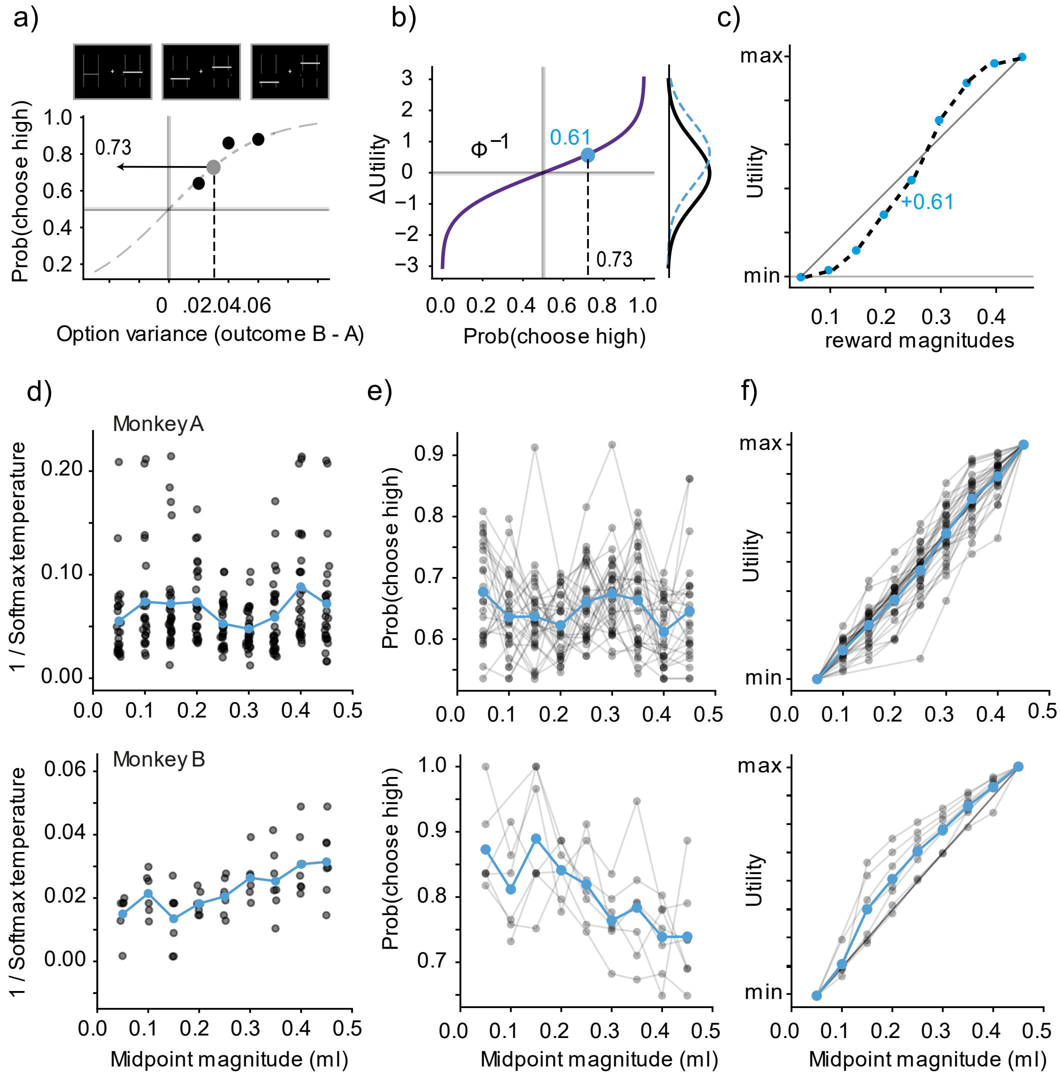


Figure 3-3 | Estimating riskless utilities from the stochasticity in safe-safe choices. a) Measuring mean stochasticity in choices between safe reward pairs. The likelihood that a monkey would pick the better of two safe rewards was measured for fixed pairs of rewards centred at different increments of the reward range. For each incremental 'midpoint', the likelihoods were fitted with softmax curves – these were then used to estimate the likelihood that the monkeys would pick the better out of two rewards if these differed by a gap of 0.03ml around the midpoint. b) Using random utility to interpret choice ratios as differences in utility. The likelihoods that monkeys would pick the better reward given a gap of 0.03ml were transformed using the inverse cumulative distribution function (CDF) of a logistic distribution. Under the random utility maximization framework, the utility of different rewards took the form of equally noisy distributions centred at the monkeys' 'true' utilities. The output of inverse CDFs is the distance between these random utilities (that is, the marginal utility at a given midpoint). c) From marginal utilities to utility. Because the distances between random utilities represented estimates of marginal utility for each midpoint, the cumulative sum of these measurements approximated a direct utility measurement for each of these midpoints. These measurements were normalized whereby the utility of the highest midpoint was 1, and the starting midpoint had a utility of 0. d) Daily strength of preference estimates from softmax fits. Each point represented the slope temperature for a softmax curve fitted on the choice ratios for three reward gaps centred on a midpoint (x-axis). The lower the inverse of the temperature parameter, the steeper the softmax curve and the more separable the random utilities were. Lower values meant higher marginal utility measurement (steeper function), higher ones meant lower marginal utility (flatter function). The blue line and points represent the mean of the daily measurements. e) Daily choice ratio estimates from softmax fits. Estimates from the same day are linked by grey lines. Ratios of 0.5 meant that the random utility distributions of the rewards separated by 0.03 ml were fully overlapping (i.e., flat utility function); choice ratios closer to 1 meant random utilities that were fully dissociated and non-overlapping. f) Normalized daily utilities. Utilities measured on the same day are linked by grey lines – these are normalized relative to the minimum and maximum midpoint fitted in the testing session. Utilities over the diagonal suggest monkeys were risk-averse, utilities under the diagonal signalled risk-seeking behaviour instead. All data in panels d-f were gathered over 22 sessions for monkey A, and 7 sessions for monkey B. Crucially, the midpoints where softmax temperatures fell above 3.5 standard deviations of the median absolute deviation were removed from the aggregate metrics and this figure (but not from the discrete choice fittings; 14 out of 265 measurements for monkey A, 3 out of 62 for monkey B).

analysis of variance (ANOVA): (i) those predicted from riskless utilities, (ii) those predicted by a merger of riskless utilities and risky probability distortion, (iii) those predicted from risky utilities and probability distortion, and (iv) the CEs actually measured on different, out-of-sample days.

3.3 Results

Prospect theory assumes that the utilities that guide both risky and riskless decisions are the same; that what differs between the two situations is simply the way that decision-makers treat probabilities (assuming the reference point is constant). I sought to validate this assumption in macaque monkeys by comparing the decisions they made in risky versus riskless choice sequences. Two rhesus macaques were trained to make binary choices between pairs of rewards presented on the left and right sides of a digital screen (fig. 3-1a). These reward options varied in terms of blackcurrant juice quantity as well as the probability that they would be delivered. The monkeys received the selected rewards after every trial – contingent on their delivery probability.

Choice preferences were elicited in trial sequences where either both options were certain and therefore riskless (probability of 100%), or in sequences where one option was safe and the other was a risky gamble between two juice magnitudes each delivered with probability $p = 0.5$ (equiprobable). I used these riskless or risky choices to infer an animal's utility function, compatible with EUT, and to predict the monkey's propensity to take risks. Both choice sequences were structured in a way that allowed me to map utilities onto aggregate behavioural metrics, and to then model these choices under the assumptions of Prospect Theory.

3.3.1 Measuring risky and riskless preferences

In risky choices, utilities were estimated by psychometrically measuring the certainty equivalent (CE) of equiprobable gambles through the fractile method: a stepwise procedure whereby one progressively sections the range of possible rewards using the CEs estimated from previous steps (Fig 3-2a,b).

Since equivariant gambles were off-limits to estimate riskless utilities, the random utility maximisation (RUM) framework – essentially an extension of EUT or PT for stochastic choices – was used to estimate changes in utility across various magnitude intervals. Under RUM, the assumption is that the utility assigned to a given reward varies randomly on every trial. That is, the subjective value of a reward is considered noisy, and varies as if it was randomly picked from a distribution of random utilities centred around the reward's 'true' utility (see this thesis' **Methods interlogue**). The precision of a monkey's choices between two rewards, under RUM, represents a direct measure of how much the random utility distributions representing the two rewards overlap.

Each riskless sequence comprised a series of choices between two safe rewards that differed by fixed juice amounts: the gap between each pair was of either 0.02ml, 0.04ml, or 0.06ml. Each of these 3 pairs were then set at unique midpoint magnitudes: 0.05ml, 0.10ml, 0.15ml, 0.20ml, 0.25ml, 0.30ml, 0.35ml, 0.40ml, or 0.45ml. The goal was to estimate the distance between the random utilities of reward pairs, for each gap, at each midpoint.

The choice ratios for each of these gap pairs, at every midpoint, were measured and fitted with psychometric curves that averaged the three ratios (minimizing the impact of any erroneous measurement on the utility estimates; Eq. 3-2). From the psychometric fits, the choice ratio at an untested gap of 0.03ml was recorded as an aggregate summary of the three measured choice ratios for each midpoint (Fig. 3-3a). These ratios were then transformed into estimates of the distance between the random utilities of the reward pairs (Fig. 3-3b); these distances represent the slope of the utility function at every given midpoint. Finally, the cumulative sum of these utility differences was calculated and normalized, providing daily utility measurements for each of the tested midpoints (Fig. 3-3c).

3.3.2 The mismatch between risky and riskless utilities

Choices measurements from risky and riskless sequences were gathered on the same day for N = 22 in monkey A and N = 7 for monkey B. For both risky and riskless sequences, the link between utility measurements and reward magnitudes was confirmed via one-way ANOVA. Looking at the effect that utility-specific gambles had on the CEs elicited from monkeys' risky choices, both monkeys exhibited a significant main effect of utility on the CEs (Monkey A: $F_{(4,123)} = 34.123$, $p = 1.838\times 10^{-24}$; Monkey B: $F_{(4,39)} = 172.537$, $p = 3.090 \times 10^{-24}$). That is, the utilities associated with different reward magnitudes were significantly different (this would not have been the case if monkeys selected options at random). In both monkeys, the utility measurements from risky choices followed an S-shape pattern: the monkeys were risk-seeking for relatively low-magnitude rewards, and risk-averse for relatively high-magnitude ones (Fig. 3-2c). For riskless choices, rather than comparing utilities, I compared the temperature of the softmax curves (i.e., the slope) that described choices at each incremental midpoint.[28] Monkey B's temperature parameters, or slopes, varied significantly depending on the midpoint magnitudes from which they had been estimated (Fig 3-3d; $F_{(8,52)} = 4.187$, $p = 6.370 \times 10^{-4}$). These did not, however, achieve significance for monkey A ($F_{(1,256)} = 1.042$, $p = 0.405$). The utility measurements for monkey B predicted risk-averse choices for the entire reward range (Fig. 3-3f, lower panel), the ones for monkey A predicted risk-neutrality (Fig. 3-3f, upper panel).

Because these aggregate measurements were not (i) Prospect-Theory-compatible, and (ii) comparable between the risky/riskless sequences[29], utilities were re-estimated and placed on a common scale using a discrete choice model of Prospect Theory (see *methods*, Eq. 3-7). In the model, utility functions took

---

[28] Since utilities were but an interpretation of the original temperature results, it was more appropriate to statistically compare the temperatures. Moreover, since these utility measurements came from a cumulative sum, they would almost always be significantly different. The proper statistical test is thus on the temperature parameter of the softmax fits – making sure that choice stochasticity did vary with midpoint magnitude.

[29] They both relied on aggregate, relative metrics and the riskless utility measurements covered a smaller range (0.05 ml to 0.45 ml) than the fractile ones (0 ml to 0.5 ml)

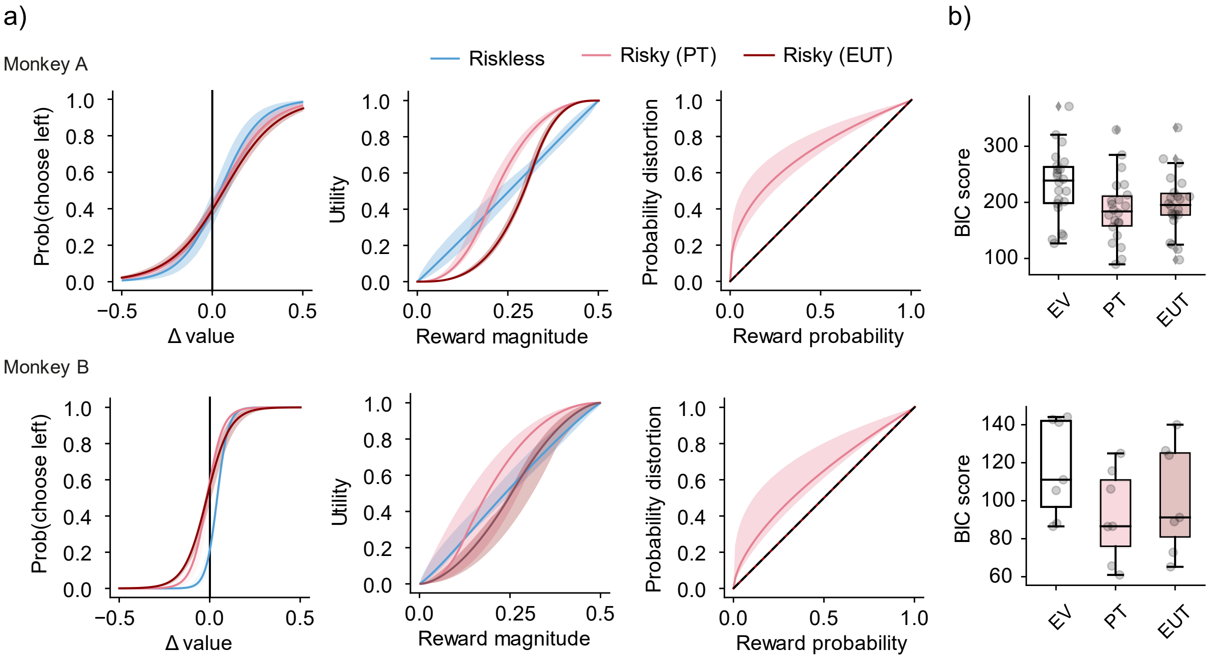


Figure 3-4 | Discrete choice estimates differ between risky and riskless choices. a) Median parametric estimates for (from left to right) the logit functions, utility functions, and probability distortion functions fitted to risky and riskless choices; the shaded area represents the 95% C.I. on the median of these functions (random sampling with replacement, n=10000).. For riskless choices, the discrete choice model predicted choices from the expected utilities of rewards (no probability distortion). Utilities were mostly linear, though slightly concave. For risky utilities, two versions of the discrete choice model were fitted: the expected utility theory (EUT) model predicted choices solely based on reward options' utilities (with no probability distortion); the other, the prospect theory (PT) model, predicted choices based on utilities and probability distortion. Both monkeys were risk-seeking; but where the PT model accounted for this mainly through probability distortion, the EUT model accounted for it through a more convex utility. b) PT fits individual choices better than EUT. Bayesian information criterions (BIC) were calculated from the log likelihoods of the daily best-fitting PT and EUT discrete choice models; an objective, expected-value based model is included for comparison.

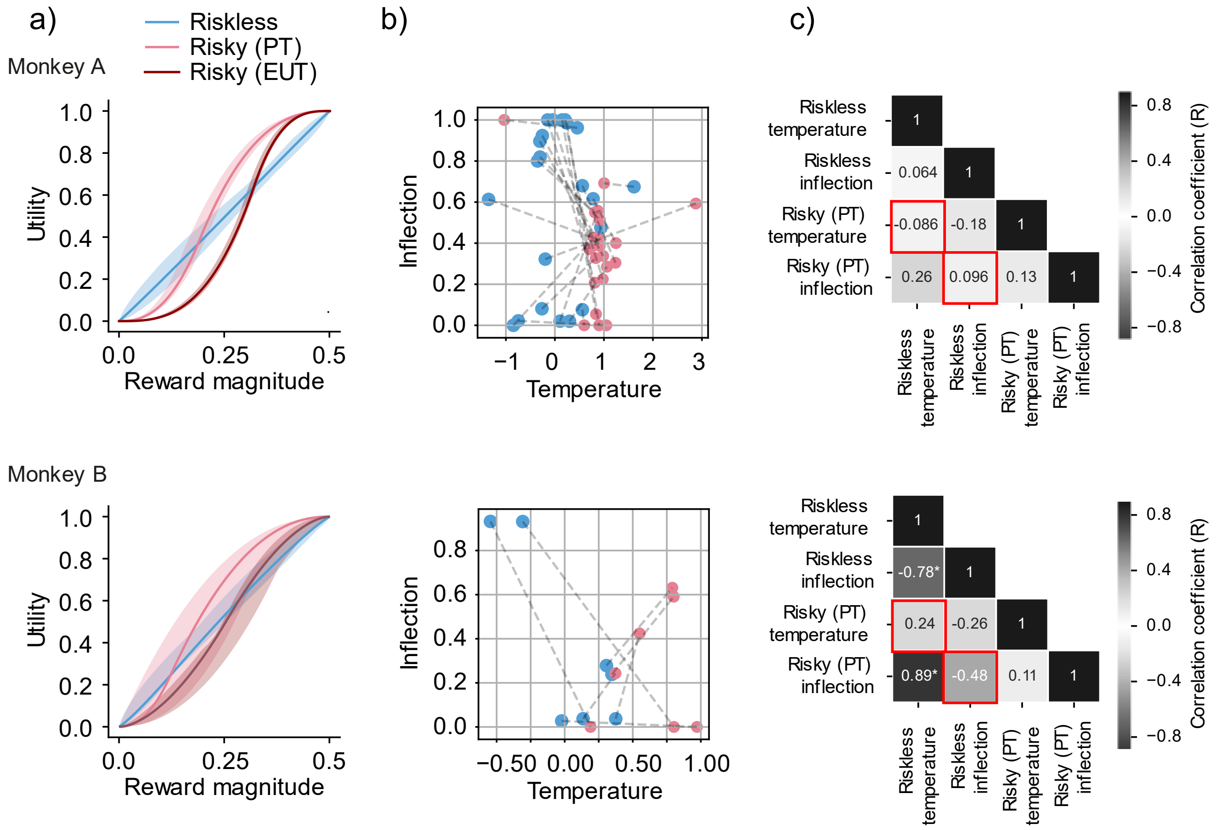


Figure 3-5 | Risky utilities do not predict riskless ones, and vice-versa. a) Median utility function estimates for risky and riskless choices. The shaded area represents the 95% C.I. on the median of these functions (random sampling with replacement, n=10000).. For riskless choices, utility estimates were mostly linear (though slightly concave). For risky utilities, the two different versions of the discrete choice model predicted S-shaped utilities, but risky EUT utility functions were more convex than PT utility functions. b) Comparison of daily parameter estimates. The sets of parameters estimated on the same day are linked via the dotted lines. The x-axis represents the log-transform of the utility functions' temperature parameters (α): $\alpha > 0$ implies an S-shaped utility function, $\alpha < 0$ implies an inverse S-shaped one. The y-axis represents the magnitude at which the utility function's curvature inverted (κ), c) Neither the temperature or inflection parameters correlate between risky and riskless utility functions. Pearson's correlations were run on the parameters from risky and riskless scenarios. Red squares highlight Pearson's R for the correlation between the temperature and inflection parameters between risky and riskless conditions. Asterisks (*) indicate significant correlations ($p < 0.05$).

the form of the cumulative distribution function of a two-sided power distribution (Eq. 3-10 ; Kotz & Dorp, 2010): a 2-parameter function that could easily account for risk-seeking and risk-averse choices (Kontek & Lewandowski, 2018). If $\alpha > 1$, the utility function would be convex and predict risk-seeking choices up to the inflection at parameter κ (predicting risk-averse choices thereafter); if instead $\alpha < 1$, the utility function would be concave and predict risk-averse behaviour up to the inflection at κ (predicting risk-seeking behaviour afterwards). For risky choices, a 1-parameter power function captured the distortion of probabilities (Eq. 3-9). Since the only probability experienced was p = 0.5, $\rho > 1$ implied an underweighting of gamble options whilst a $\rho < 1$ implied that gambles were overweighed instead. Three forms of this discrete choice model were fit onto daily experimental results (Fig. 3-4a). For riskless choices, only utility's parameters were free (probability distortion was fixed as linear); for risky choices, one version of the model incorporates both a utility function and a probability distortion function (PT utilities), the other versions relied solely on a utility function (EUT utilities).

In terms of risky choices, both the EUT and PT models predicted s-shaped utility function. The PT model, however relied on a concave probability distortions (significantly so; one-sample t test, Monkey A: $t_{(21)} = -4.912$, $p = 7.383 \text{ X } 10^{-5}$; Monkey B: $t_{(6)} = -4.196$, $p = 0.006$), rather than convex utility function, to predict risk-seeking behaviour. For that reason, PT's s-shaped utility functions were mostly left-skewed (more concave than convex) whereas EUT utility functions captured risk-seeking behaviour solely through a right-skewed s-shape (more convex than concave). Overall, the daily best-fitting parameters from the PT and EUT models were significantly different from each other (Table. 3-1), with the PT model capturing behaviour significantly more reliably (Fig. 3-4b; Wilcoxon rank sum test; monkey A: $p = 2.00 \times 10^{-4}$; monkey B: p = 0.018).

Like in previous human studies (Abdellaoui et al., 2007; Stalmeier & Bezembinder, 1999), estimating utilities through PT's framework, rather than EUT's one, brought risky fits more in line with riskless ones (Table. 3-1; Fig 3-5a). However, one-way MANOVA analyses on the risky PT and riskless parameters (including noise and side-bias) revealed that significant differences between the two types of choices were maintained (Fig. 5b). Looking specifically at utility, both monkey A and B's utility temperatures were consistently and significantly different between risky and riskless choices (Monkey A: $F_{(1,42)} = 6.078$, p = 0.018; Monkey B: $F_{(1,12)} = 13.883$, p = 0.003). Simply put, the monkey's risk-attitudes were significantly different. Independently, monkey B's noise and side bias parameters were significantly different across risky and riskless choices (noise: Monkey A: $F_{(1,42)} = 1.838$, p = 0.182; Monkey B: $F_{(1,12)} = 9.562$, p = 0.009; side bias: Monkey A: $F_{(1,42)} = 0.021$, p = 0.885; Monkey B: $F_{(1,12)} = 17.889$, p = 0.001), and Monkey A saw a significant difference between the position of utility's inflection (Monkey A: $F_{(1,42)} = 5.277$, p = 0.023; Monkey B: $F_{(1,12)} = 0.204$, p = 0.659). For monkey A, this meant that riskless utilities predicted mostly one set of risk-attitudes, whilst risky utilities predicted two (risk-seeking to risk-averse, depending on the magnitudes involved). For monkey B, a difference on both the noise and side bias parameter reflected a difference in the way the monkey guided his choices. In addition, there was no clear relationship between the parameters of risky utility functions and riskless ones (Fig. 3-5c).

3.3.3 Prospect theory predictions are task-specific

To extend these findings beyond parametric comparisons, and to understand how these parameters translated into predictable behaviour, the risky and riskless utility estimates (and probability distortion) were used to predict the certainty equivalents of different out-of-sample gambles. Four equiprobable ($p = 0.5$) gambles were defined (Fig. 3-6a), and their CEs were derived from the daily fitted utility functions (risky PT and riskless) and probability distortion functions (risky PT). In addition to risky- and riskless-derived CEs, and since riskless parameter relied on no probability distortion, a third set of CEs was derived from a hybrid model of riskless utilities and risky probability distortions measured on the same day (risky/riskless hybrid). A total of twelve CE values were estimated for each testing day: four stemming from risky estimates, four from riskless ones, and four from the risky/riskless hybrid.

Essentially, each set of predicted CEs served as a testable hypothesis regarding the relative value of gambles versus safe reward options. I therefore presented monkeys with sequences of out-of-sample trials to measure the true CEs of these gambles and compare these to the predictions made from the parametric utility fits. On a separate set of days (17 days for monkey A, 29 for monkey B), I presented the monkeys with the four predefined gambles pitted against safe rewards (varying by 0.05ml), and estimated their CEs using the earlier psychometric curve (Eq. 3-1). The CEs that had been directly measured were compared to those predicted from the risky and risky/riskless hybrid models using multiple two-way ANOVAs. The purely riskless predictions were not included in the ANOVA analyses since, (i) the CEs were directly correlated with the risky/riskless hybrid CEs, and (ii) they did not incorporate the effects of a probability distortion, which had previously been identified as contributing to a significantly more reliable fit for risky choices. In both monkeys, there were significant difference between all pairings of the various CE distributions (Table. 3-2; $p < 0.05$).

Finally, to understand where the true CEs differed, in terms of utility and probability distortion, I fit the previously used discrete choice model to the choices made by the monkeys in the CE elicitation sequence (Fig. 3-7). These new sets of decision parameters were compared the those from the 'fractile' risky estimates. As predicted by the previous CE estimates, one-way MANOVA analysis (this time including probability distortions) revealed that the parameters fit to the latter CE estimation sequences were significantly different to the parameters that fit monkey A's choices during fractile sequences (Monkey A: $F_{(1,37)} = 12.412$, Wilks's $\lambda = 0.347$, $p = 8.093 \times 10^{-7}$). That is, the fits describing choices in two risky scenarios, both using the exact same stimuli, led to remarkably different decision parameters (noise: $F_{(1,37)} = 0.114$, $p = 0.737$; side bias: $F_{(1,37)} = 40.01$, $p = 2.28 \times 10^{-7}$; utility temperature: $F_{(1,37)} = 1.819$, $p = 0.186$; and utility inflection: $F_{(1,37)} = 12.901$, $p = 9.499 \times 10^{-4}$) – in particular the probability distortion ($F_{(1,37)} = 14.678$, $p = 4.776 \times 10^{-4}$). Surprisingly, however, while monkey B's risky CE

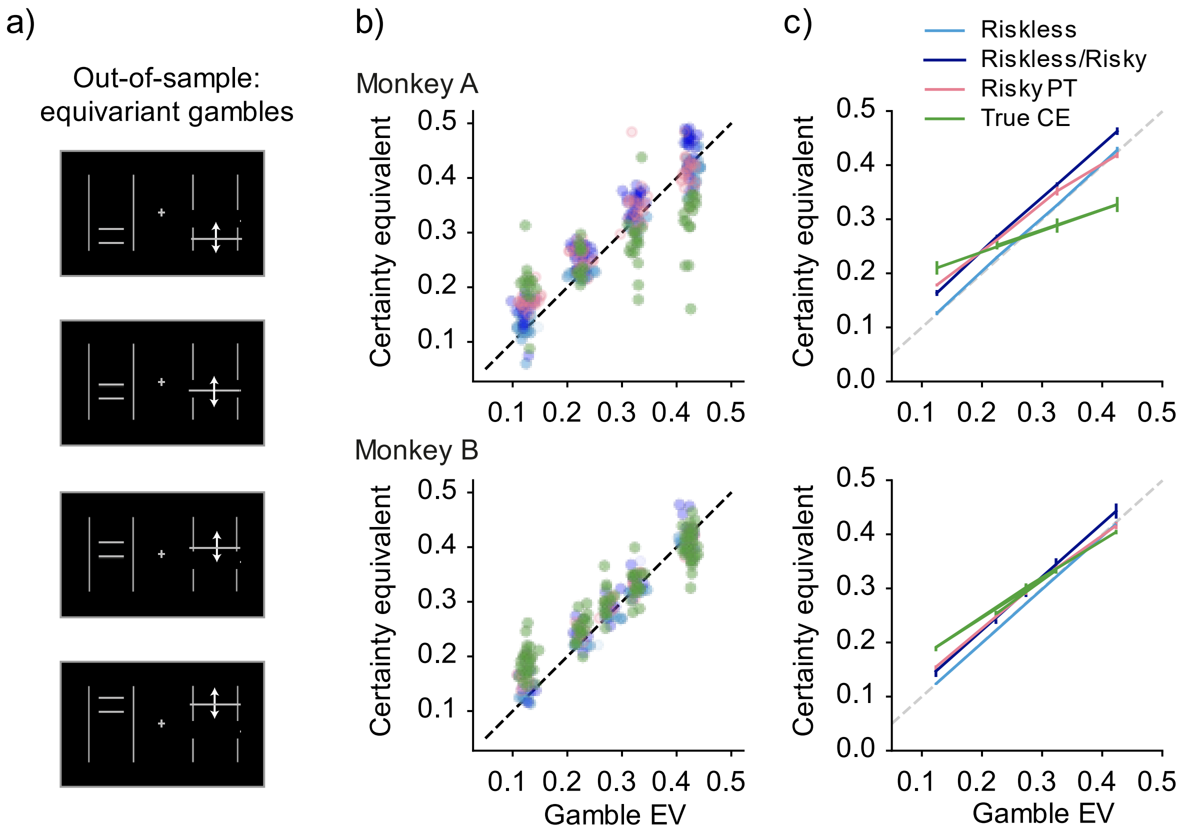


Figure 3-6 | Out-of-sample comparison of predicted and measured certainty equivalents (CEs). a) the CEs of four different gambles were measured. Each of these gambles had a magnitude spread of 0.15ml, and their expected values were anchored at 0.125ml, 0.225ml, 0.325ml, and 0.425ml. b) Distribution of measured CEs and predicted ones. Each point represents the predicted CEs from the utility parameter estimates on a given day; the green ones are real CE measurements from monkeys' choices. There were, for monkey A, 62 measured CEs and 88 predicted ones; for monkey B, there were 124 measured CEs and 35 predicted ones. c) Average distribution of CEs. The average CEs, in addition to the raw CEs, is plotted for clarity. Measured CEs differed significantly from all predicted one. Both monkeys were a lot more risk-seeking (for low reward magnitudes) in risky sequences where risk was kept constant than for risky sequences where risk varied (fractile method).

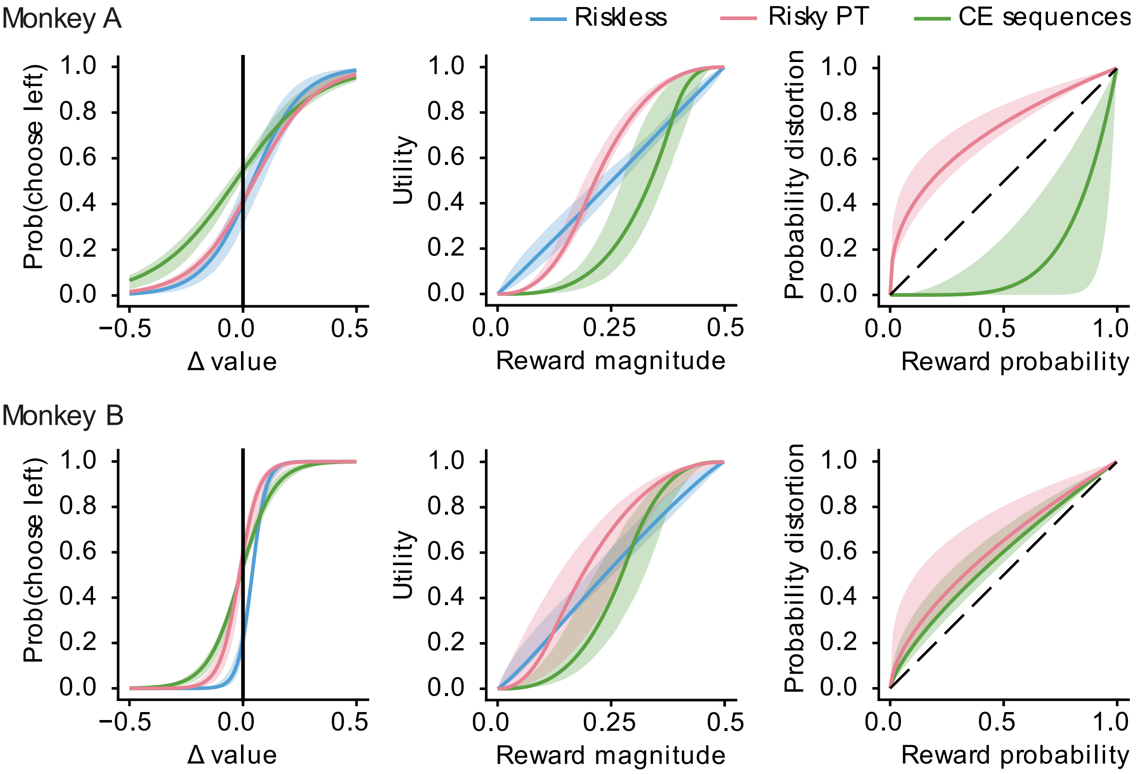


Figure 3-7 | Comparing decision parameters in riskless, risky equivariant, and risky non-equivariant choices. Median parametric estimates for (from left to right) the logit functions, utility functions, and probability distortion functions fitted to risky and riskless choices; the shaded area represents the 95% C.I. on the median of these functions. Most noticeably, monkey A appeared to treat uncertainty very differently depending on the sequence trials used to elicit risky decision parameters. The PT utility functions estimated from both monkeys' out-of-sample CE elicitation sequences were much more convex than those estimated from sequences used to build utilities in fractiles.

predictions and measured CEs differed significantly, the difference between the parametric fits on his individual choices failed to reach significance (Monkey B: $F_{(1,34)} = 1.544$, Wilk's $\lambda = 0.795$, $p = 0.206$). None of the decision parameters were significantly different (noise: $F_{(1,34)} = 0.086$, $p = 0.361$; side bias: $F_{(1,34)} = 0.025$ $p = 0.876$; utility temperature: $F_{(1,34)} = 1.098$, $p = 0.302$; utility inflection: $F_{(1,34)} = 3.547$, $p = 0.068$; and probability distortion: $F_{(1,34)} = 0.422$, $p = 0.521$).

3.4 Discussion

Using a robust, incentive-compatible task, we showed that utility functions that describe decisions involving risk more closely mimicked riskless utility functions if probability distortions were considered. I modelled macaque monkeys' risky and riskless choices through stochastic versions of PT and EUT, and reliably estimated functional parameters that best described their choices. Each day, the monkeys were presented with risky or riskless binary choice sequences. In risky ones, they made choices between gambles and safe rewards; in riskless ones, both choices had a single, certain outcome. I found that modelling monkeys' risky decisions via the PT model of choice, in addition to providing a better fit than EUT, led to decision parameters that more closely resembled riskless ones. This trend is in-line with the human literature (Abdellaoui et al., 2007; Stalmeier & Bezembinder, 1999). However, the final comparison differed; the monkeys' utility functions were more alike, but they were still significantly different in terms of shape and position.

In terms of behavioural metrics, the CEs estimated in fractile sequences suggested both monkeys were risk-seeking for all but the highest of reward magnitudes that they experienced. The PT and EUT models predicted similar risk-seeking behaviour via an overweighing of gamble options, but they differed in the way in which they achieved this. Both EUT and PT models predicted s-shaped utilities, the PT model, however, accounted for the monkey's risk-seeking behaviour mostly through its concave probability distortion. In other words, the subjective probability of 'winning' a gamble was higher than the objective probability of winning regardless of utility's effects. EUT fits on the other hand, captured risk-seeking behaviour exclusively through their utility function; one that was right-leaning (more convex than concave) and so predicted higher utilities for gamble options than for safe ones. Since PT's utilities were 'free' from the effects of probability distortion, the s-curves were left-shifted (i.e., more concave than convex), suggesting a relatively more risk-averse utility function than from EUT's predictions. Comparing these findings to the riskless utility fits, we found that PT utilities deviated far less from riskless utilities than EUT ones. Still, the utilities estimated from riskless binary choices were relatively linear (if slightly risk-averse), a shape that was at odds with that of the risky PT estimates. It appears that, at least within the confines of our experiment, the difference between risky and riskless utilities was not as simple as the addition of a probability distortion parameter. That is, the monkeys likely chose using different strategies in the two conditions: where risk significantly bias-distorted the way they evaluated options, both monkeys appeared to be far more objective in their preferences for riskless rewards.

Of course, while the same binary choice design was used in risky and riskless choices, the difference between options was much greater in risky sequences than in riskless ones. To estimate aggregate riskless utilities, for example, the rewards that the monkeys experienced differed only by up to 0.06ml in every trial. In risky sequences, on the other hand, gambles were pitted against safe rewards spread over the full range of the gambles' outcomes. Monkeys experienced a broad range of magnitudes in each of the sequences, but the differences between riskless choices could have required far more attention to dissociate than those in riskless choices (something we cannot account for; but see, Farashahi et al., 2018). The logical next step will be to see if risky choices between closer-together gambles reproduce the patterns of choices that I have observed.

Where these findings fail to replicate the data from risky and riskless introspective studies (though see Hertwig, Wulff, & Mata, 2018), they are nonetheless in line with the incentive-compatible time trade-off approach. Since these types of time discounting tasks are easily adapted to study preferences in rhesus macaques (Blanchard et al., 2013; Hayden & Platt, 2007; Hwang et al., 2009; S. Kobayashi & Schultz, 2008), it would be interesting to see how utility functions estimated using time trade-offs in macaque monkeys correlate with the present findings. Another approach that would be interesting to consider is the one used by Chung, Glimcher and Tymula (Chung et al., 2019), where they compared risky and riskless choices between bundles of outcomes - estimating utilities through identifying the combinations of rewards for which decision-makers are indifferent. They found that risky and riskless choices could be reconciled when choices involved gains, but that PT failed to reconcile the two when the choices involved losses. Since preferences over losses are generally risk-seeking (for humans), it could be that the macaque monkeys' risk-seeking behaviour mimics this loss-related discrepancy. If macaque monkeys were to, in risky settings, adjust their expectations in a way that paints the lower outcome of a gamble as a loss, one would expect the lower end of their utility function to behave like the loss side of PT's value function (Kahneman & Tversky, 1979). There is some evidence that rhesus macaques (and indeed humans) do this: they exhibit preferences consistent with win-stay lose-shift strategies (G. Barron & Erev, 2003; Gilovich et al., 1985; Heilbronner & Hayden, 2013). For repeated gamble-safe choices, they generally reverse their risk-seeking preferences for gambles depending on if they have previously won or lost a previous gamble instance (Blanchard, Wilke, et al., 2014; Ferrari-Toniolo et al., 2019; Lau & Glimcher, 2005). If this is the case, fitting macaques' choices through utility models that account for trial-by-trial changes in preferences are likely to do a better job at reconciling risky and riskless utilities than using a single utility and probability distortion function applied to the entire experimental procedure.

Overall, the results presented here add to the already dire need for decision models to account for flexible, context-specific preferences (Farashahi et al., 2018; Hayden et al., 2008; Heilbronner & Hayden, 2016). Although I expected some form of discrepancy between risky and riskless utilities, the fact that one of the monkeys' probability distortions also shifted significantly between risky scenarios (equivariant CE elicitation and fractile method) suggests that preferences flexibly adjust to more than just the level of risk at hand. For decision theory as a whole, reconciling dynamic preferences with more traditional economic models would go a long way to making more accurate, descriptive predictions.

Table 3-1 | MANOVA Tests for pairwise differences between the risky EUT, risky PT, and riskless discrete choice models. The analyses were run on four of the five free parameters, excluding the probability distortion one. The risky EUT and riskless models had no probability distortion parameter to compare with the risky PT model's probability distortion one.

	Utility Type	F (1, 42)	p	Wilks λ
Monkey A	Riskless, Risky (PT)	3.757	0.011	0.722
	Risky (EUT), Risky (PT)	8.625	4.364×10^{-5}	0.531
	Riskless, Risky (EUT)	4.475	0.005	0.685
	Utility Type	F (1, 12)	p	Wilks λ
Monkey B	Riskless, Risky (PT)	8.649	0004	0.206
	Risky (EUT), Risky (PT)	2.371	0.130	0.487
	Riskless, Risky (EUT)	8.719	0.004	0.205

Table 3-2 | Two-way ANOVA Tests for pairwise differences between three sets of certainty equivalents. The certainty equivalents came from the daily predictions of the risky/riskless hybrid model, the PT model, and the ones measured from out-of-sample sequences.

	Utility Type	Df	F	p
Monkey A	Risky/Riskless, Risky (PT)	(7, 168)	404.058	1.384×10^{-101}
	Risky/Riskless, True CEs	(7, 142)	134.906	1.746×10^{-59}
	Risky (PT), True CEs	(7, 142)	88.538	1.305×10^{-48}
	CE Type	Df	F	p
Monkey B	Risky/Riskless, Risky (PT)	(9, 60)	171.021	2.426×10^{-39}
	Risky/Riskless, True CEs	(9, 149)	202.183	7.959×10^{-79}
	Risky (PT), True CEs	(9, 149)	211.873	3.189×10^{-80}

Chapter 4

Choice sequence shapes probability distortion in rhesus macaques

4.1 Introduction

Choices between uncertain rewards require decision-makers to evaluate each option along multiple dimensions. At the very least, a decision-maker needs to simultaneously consider the reward's quantity and probability of occurrence if he is to evaluate its attractiveness in relation to other choice prospects. The von Neumann and Morgenstern Expected Utility (EU) theory was the first axiomatic model of rational behaviour capable of describing people's choices in these situations (von Neumann & Morgenstern, 1944). EU theory rigorously introduced the concept of utility as a representation of a decision-maker's subjective value for an objective reward quantity. Through the metric of utility, EU theory was able to describe different risk attitudes, such as the risk-seeking behaviour of a gambler or the risk aversion of an insurance buyer; it was, however, soon challenged by the various experimental results of behavioural economics (for review see, e.g., Machina, 1987; Starmer, 2000; Weber & Camerer, 1987). Attempts to resolve some of these challenges led to the development of several generalized expected utility theories, many of which (notably prospect theory, rank-dependent utility theory, and cumulative prospect theory) incorporated the concept of probability distortion (Kahneman & Tversky, 1979; Quiggin, 1982; Tversky & Kahneman, 1992). While maintaining the nonlinear relationship between subjective utility and objective reward magnitudes, these theories made use of subjective probability weightings, or probability distortions, to account for the idea that reward probabilities were also treated nonlinearly during choice.

Experimental measures of probability distortion in humans and monkeys typically show that, whereas small probabilities tend to be overweighed by decision-makers, large probabilities are instead underweighed (Gonzalez & Wu, 1999; Kahneman & Tversky, 1979; Stauffer et al., 2015). There is, however, dramatic variation in this pattern of distortion across both different subjects (Bruhin et al., 2010; Burke et al., 2018; Gonzalez & Wu, 1999) and between different task contexts (Farashahi et al., 2018; Hertwig et al., 2004; Wu et al., 2009). While the causes of such variability have yet to be identified, differences in probability distortions could relate to the way in which probability information is presented to decision-makers (Hertwig et al., 2004), or the way in which probability knowledge is acquired and stored by the decision-maker (Camilleri & Newell, 2013). Some studies suggested that prospect theory might altogether be incapable of explaining differences in risk attitudes across these contexts (Kellen et al., 2016).

Here we investigated the role of choice context, specifically sequence structure, as a possible source of probability distortion variability in rhesus macaques, animals known to show quantifiable and reproducible probability distortions (Stauffer et al., 2015). To achieve this, we first measured the certainty equivalents (CEs) of specific gambles, defined as the amount of reward for which the animal was choice-indifferent with regards to said gambles; the CE therefore indicated the subjective value of the gamble in the 'currency' of the safe reward. We then simultaneously estimated the contributions of utility and probability distortion to these subjective values, allowing us to model the shape of the monkeys' probability distortion independently from utility.

We used this technique to investigate the possible influence of trial sequence structure on the shape of the probability distortion in two different task situations: randomly intermixing the trials required for the CE measurements of all gambles, or determining the CEs of different gambles via separate blocks of trials. We performed an out-of-sample test to validate and extend the results of our main task, and investigated the contribution of trial history as a possible correlate of probability distortion variance. Our data showed that a change in the presentation order of probability information indeed altered the observed probability distortion pattern, inducing a reversal in probability distortion shape.

4.2 Methods

4.2.1 Animals and experimental setup

Two male rhesus macaques (Macaca mulatta) were used in this study (11.2 and 13.2 kg). During experiments, the monkey sat in a primate chair (Crist Instruments) and made choices between two rewarding stimuli presented on a computer monitor positioned 30 cm in front of them. The animals reported their choices between options with a left-right motion joystick (Biotronix Workshop, the University of Cambridge). Joystick position and task event times were sampled and stored at 1 kHz on a Windows 7 computer running custom-made software written in MATLAB (The MathWorks, 2015a) using Psychtoolbox (version 3.0.11). All experimental protocols were assessed and approved by the Home Office of the United Kingdom.

4.2.2 Experimental design

We trained the monkeys to associate visual stimuli with specific juice rewards that varied along two dimensions: the quantity of juice delivered (reward magnitude, m) and the delivery probability of the reward (reward probability, p). To capture both dimensions descriptively, the visual stimuli consisted of a horizontal bar or of a pair of horizontal bars framed between two vertical lines. The vertical position of the horizontal bars signalled the magnitude of juice delivered; the width of the bar signalled the probability of their delivery from no bar (no reward) to touching the frame on both side (certain reward). To ensure that the bar's edge position relative to the frame was not used as a cue for the gamble's mathematical expected value (EV; i.e., the product of m and p), the bars were randomly shifted horizontally on each trial. This guaranteed that magnitude and probability information were independently presented and used to make choices. Multiple partial bars found between the vertical frames signalled gambles between 'risky' rewards, whereas a singular, full-width horizontal bar signalled a safe, riskless reward. Across all trials, monkeys experienced rewards ranging from 0 ml to 0.5 ml in 0.05 ml increments, and gamble probabilities varying between 0.1 and 1 in decimal increments (0.1).

The animals learned to associate rewards and magnitudes with the visual stimuli schema through >10,000 single-outcome, or imperative, trials. In these trials, only one option was presented on either side of the screen. To obtain the cued reward, the animals were required to select the side on which the reward was presented. All reward options were repeated on both the left and right sides of the computer screen, alternating pseudorandomly to control for any side preference.

Following imperative training, we presented the animals with a binary choice paradigm where they had to choose one of two reward options presented simultaneously. Most binary choice trials pitted a safe reward against a gamble. All gambles consisted of two probabilistic rewards: the monkey could either get a fixed 0.5 ml of juice with probability p, or 0ml of juice with probability $1 \geq p$. Safe options varied in terms of reward magnitude only. In separate sets of trials, we presented the animals with choices between two gambles with two outcomes each (possible outcomes: 0, 0.25, 0.5 ml). In these trials, one of the gambles could have two non-zero outcomes (0.25 and 0.5 ml). In all cases, reward was delivered probabilistically, matching the probabilities cued by each stimulus.

Trials began with a white cross at the centre of a black screen, followed by the appearance of a joystick-driven cursor. The cursor had to be moved to the centre cross in order for a trial to begin. After successfully maintaining the cursor on the central cross for 0.5–1 s, two visual option cues appeared left and right of the central cross (see Fig. 4-1a). In the case of imperative trials, only one option appeared while the other side remained dark. The animal had 3 s to convey his decision by moving the joystick to the selected side, after which the unselected option would disappear. The animal's response time (RT; i.e., the time interval between the cues appearance and the beginning of the joystick movement) was collected for individual trials. Reward delivery occurred after the holding time (0.1–0.2s), and the selected option lingered on the screen for 1s post reward delivery to reinforce stimulus-reward associations with visual feedback. A variable intertrial period of 1–1.5s (blank screen) led to the next trial onset. Unsuccessful central hold, side selection hold, or trials where no choices were made resulted in a 6s timeout for the animal, after which the trial would be repeated.

4.2.3 Psychometric elicitation of CEs

The likelihood of a monkey choosing a specific, individual gamble over different safe options was assessed through the binary choice paradigm (see Fig. 4-1b). The resulting choice ratios were then used

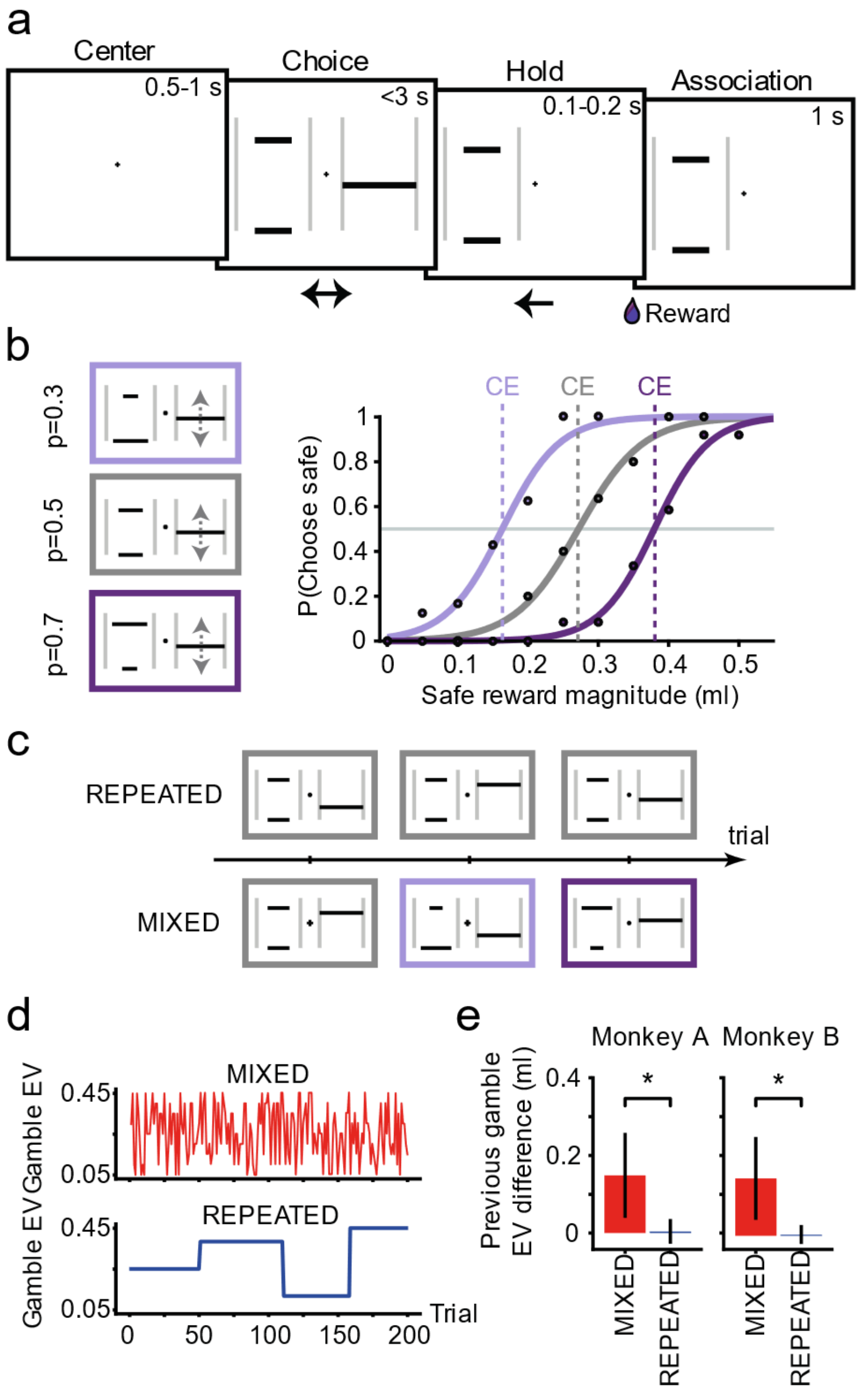


Figure 4-1 | Experimental design. a, Trial sequence. Each trial started with the monkey moving a white cursor, through left/right arm movements with a joystick, to the centre of the screen. After 0.5–1 s (centre holding), two cues appeared indicating the two offered options (choice period): possible reward magnitudes and probabilities were indicated by the vertical position and width of a horizontal bar, respectively. A single horizontal bar represented a sure reward. Two bars represented a gamble with two possible outcomes. The monkey moved the cursor to the side of the preferred option, within 2 s. After 0.1–0.2 s (holding time), the juice reward was delivered according to the chosen option's reward magnitude and probability. A further 1 s (association period) followed to reinforce the association between chosen cue and reward. b, Psychometric elicitation of CEs. Left, Three example gambles with different reward probabilities (p = 0.3, p = 0.5, p = 0.7) paired with varying safe magnitudes to elicit each gamble's CE. Right, Each point represents the probability of choosing the safe option in choices between a fixed gamble (identified by the colour) and a varying safe magnitude (horizontal axis). Lines are psychometric curves obtained by fitting a softmax function to the choice ratios. Each line is associated with one specific gamble and identifies its CE as the magnitude corresponding to a choice ratio of 0.5 (vertical dashed line). c, Task conditions. The CEs were elicited using two sequence structures: in the MIXED condition, different gambles and different safe options were randomly intermixed; whereas in the REPEATED condition, the CE measurement for one gamble was completed before presenting a different gamble. d, Temporal sequence of the presented gamble EV in the two elicitation conditions for one sample session (first 200 trials). The trial-by-trial variation of the gamble EV highlights the difference in sequence structure between MIXED (red) and REPEATED (blue) conditions. e, Variability of gamble EV across consecutive trials. Absolute value of the gamble EV difference (mean ± SEM) between two consecutive trials, showing the main distinction between the two elicitation sequences: the previous trials' gamble EV was consistently different from the current one in the MIXED condition, whereas it stayed constant in the REPEATED condition. *Significant difference between conditions (t test).

to fit a logistic sigmoid function, or psychometric curve, to estimate choice likelihoods for every possible safe-gamble pairing within the tested reward range as follows:

$$P_{(ChooseSafe)} = 1/\left(1 + e^{-\left(\frac{SafeReward_{ml} - x_0}{\sigma}\right)}\right) \qquad (4\text{-}1)$$

These psychometric curves captured the likelihood of choosing a safe option over a gamble through two free parameters: x_0, measuring the x-position of the curve's inflexion point, and σ, the function's temperature parameter, reflecting the steepness of the curve. Importantly, only sequences that contained choices between a gamble and a minimum of three different safe options (repeated at least 4 times) were used in the analysis.

The point of choice indifference between gamble and safe options, corresponding to the inflexion point x_0 of the resulting model, represented a gamble's CE: the certain safe reward that was of equal subjective value to the gamble. CEs could then be used to categorize behaviour. Gambles where the CEs were of greater value than the predicted EV signalled risk-seeking behaviour for that gamble's probability value. Gambles with CEs lower than their EVs indicated risk-averse behaviour for that option. For cases where CEs were equal to EVs, the animals were seen as being risk-neutral.

To explore the role of task structure on the variability of one's probability distortion pattern, we measured CEs in one of two elicitation conditions: MIXED or REPEATED trial sequences (see Fig. 4-1c,d,e). In the case of MIXED sequences, multiple CEs were elicited through single blocks of randomized choice trials involving different gambles and safe options. Such blocks were repeated until each gamble-safe pair had been presented a minimum of four times each. In the case of REPEATED sequences, CEs were elicited using blocks of trials that contained a unique gamble. These REPEATED trial blocks pitted multiple safe options against a single gamble for the elicitation sequence. Other than these sequence designs, everything from visual cues to timescales was identical. The only difference between elicitation conditions was the number of different probabilities of reward (gambles) experienced within a trial block. Testing for each elicitation condition was done consecutively over multiple days, with each monkey receiving imperative training before their daily elicitation sessions. We collected on average 172.95 ± 20.24 (mean ± SEM) trials per daily session over 56 sessions for Monkey A (22 REPEATED and 34 MIXED sessions, in consecutive days), and 414.63 ± 27.87 trials over 59 sessions for Monkey B (31 REPEATED and 28 MIXED sessions, in consecutive days).

4.2.4 Analysis of behavioural data

All data were collected, stored, and analyzed using custom MATLAB and Python (Python 3.7.1, SciPy 1.1.0; Oliphant, 2007) software. Analyses were run on trial-by-trial choice data, and on the CEs elicited psychometrically from these trial-by-trial choices. The data were stored and analyzed separately for the two animals.

Before any comparative analyses, the use of visual stimuli to guide the monkeys' decision behaviour was verified through analysing all CE elicitation trials (excluding error trials where the animals made no choices) in a logistic regression model as follows:

$$y = \beta_0 + \beta_1(EV_{Gamble}) + \beta_2(EV_{Safe}) + \beta_4(Risk) + \beta_3(Position_{LR}) + \varepsilon \qquad (4\text{-}2)$$

The dependent variable took a value of y = 1 if the gamble was chosen and y = 0 if the safe option was chosen instead. As had been previously done (Stauffer et al., 2015), we fitted four independent variables: option values (EV_{gamble}, EV_{safe}) were defined as the EVs of gamble and safe rewards; gamble position ($Position_{LR}$) as 0 for left, 1 for right screen side; and the outcome's risk value (Risk) was defined as $\sqrt{p * (1-p)}$, a proportional representation of probabilistic variance. We fitted individual testing days separately, fully standardizing the β-coefficients and then testing for statistical significance (one-sample t test, $p<0.05$) to identify relevant decision variables. Positive regression coefficients indicated an increase in the likelihood of choosing a gamble over a safe option with increasing independent variable value; negative regression coefficients indicated a decrease in the likelihood of choosing the gamble.

Once the use of onscreen stimuli to guide choices had been confirmed, CEs were measured using the aforementioned psychometric fit. CEs gathered in the MIXED condition were compared with CEs gathered under the REPEATED condition using a two-factor ANOVA with gamble probability and elicitation condition as main factors. The ANOVA also captured any interaction between the two factors, highlighting any condition effects present at a sequence level.

We used trial-by-trial choices to parametrically model the respective effects of utility and probability distortion on single choices, and more generally, on the subjective value of gambles (CEs). For each daily testing session, we simultaneously estimated both the utility and probability distortion functions from within a standard discrete choice model. Functional parameters that best-described choices between gamble-safe pairs were elicited in this way, capturing the individual effects of nonlinear utility and probability distortion. The model ran on trial-by-trial choice data, with data binned into several sets containing one gamble and all safe options presented against it on the day (CE elicitation sequence). The discrete choice (softmax) function returned the probability of choosing the gamble option based on the subjective value of both the gamble (V_G) and the safe reward presented (V_S).

$$P_{choose\ Gamble} = 1/(1 + e^{-\lambda(V_G - V_S)}) \qquad (4\text{-}3)$$

The softmax parameter, λ, defined the likeliness of choosing the better prospect; each option's value (V) was defined according to prospect theory (Kahneman and Tversky, 1979), as the product of utility (u) and probability distortion (w) outputs as follows:

$$V(p,m) = w(p) \times u(m) \qquad (4\text{-}4)$$

Utility was modelled through a power function as follows:

$$u(m) = \left(\frac{m_{outcome}}{m_{max}}\right)^{\rho} \tag{4-5}$$

where ρ>1 captured risk-seeking choice behaviour, ρ<1 captured risk-averse choice behaviour (ρ<1), and p = 0 implied risk neutrality (Hsu et al., 2009). Magnitude values were divided by 0.5 ml (m_{max}), such that the maximal reward a monkey could get was anchored at 1 unit of utility.

We compared four functional models of probability distortion in an attempt to best capture changes in probability distortion across conditions. Of these classical models, two had a single fitting parameter: the one-parameter Prelec function (Eq. 4-6, Prelec-1, parameter: α) and the Kahneman and Tversky probability weighting function (Eq. 4-7, Tversky, parameter: ε); the others had two fitting parameters: the two-parameter Prelec function (Eq. 4-8, Prelec-2, parameters: α, β) and the Gonzalez and Wu log-odds model (Eq. 4-9, Gonzalez, parameters: γ, δ). Formally:

$$w(p) = e^{-(-\ln(p))^{\alpha}} \tag{4-6}$$

$$w(p) = \frac{p^{\varepsilon}}{(p^{\varepsilon}+(1-p)^{\varepsilon})^{1/\varepsilon}} \tag{4-7}$$

$$w(p) = e^{-\beta(-\ln(p))^{\alpha}} \tag{4-8}$$

$$w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma}+(1-p)^{\gamma}} \tag{4-9}$$

Using a maximum likelihood estimation (MLE) method, we simultaneously estimated the functional parameters (Θ) from the experimental data. We defined the log-likelihood function as follows:

$$LL(\Theta|y) = \sum_{i=1}^{n} y_i \times log(P_{Choose\ Gamble}) + \sum_{i=1}^{n} y_i' \times log\left(P_{Choose\ Safe}\right) \tag{4-10}$$

The log-likelihood function was defined on all trials in a session (n), the trial number (i), and the choice outcome variable for the gambles and safe options (y and y', respectively). The outcome variables took a value of 1 if their respective option was chosen; 0 otherwise. We used an unconstrained Nelder–Mead search algorithm (Python: scipy.optimize.minimise) to compute the functional parameters that minimised the negative log-likelihood (-LL). This MLE approach allowed for the simultaneous estimation of the model's free parameters, placing no constraints on their values (Abdellaoui, 2000; Pelé et al., 2014; Stauffer et al., 2015).

The algorithm identified the best fitting softmax, utility, and probability distortion parameters with respect to each monkey's daily choices in CE elicitation sequences. Four complete models were parametrized, accounting for the different probability distortion functions investigated. From these, we calculated the Bayesian Information Criterion (BIC) to pinpoint the probability distortion function most reliable in capturing behaviour. Four sets of parameters and their BIC were estimated for every testing day, independently for each model. We selected a single model for further analysis, based on the flexibility of the functional model, its comparative BIC score (one-factor ANOVA with repeated measures, Greenhouse-Geisser–corrected p values: pGGc), and the deviance between the model's predicted CEs and the experimental ones (one-factor ANOVA with repeated measures, Greenhouse-Geisser–corrected p values; Greenhouse & Geisser, 1959).

We further validated the parameter estimation procedure by running 10 simulated choice datasets within the fitting algorithm. Datasets used for testing were generated by fixing the utility parameter (ρ) and varying the probability distortion parameter (α), or vice versa. The softmax temperature parameter was kept constant ($\lambda = 10$), as we specifically wanted to test the robustness of the estimation procedure in relation to variability in the utility and probability parameters. These fixed models were used to simulate individual trial choices. We simulated 6 trials for every gamble-safe pair (safe magnitude levels: 0–0.5 ml in steps of 0.05 ml). Five datasets varied in terms of utility ($\rho = 0.20, 0.50, 1.00, 1.50, 3.00$), and five in terms of probability distortion ($\alpha = 0.33, 0.67, 1.00, 1.50, 3.00$). We measured estimation accuracy as the 95% CI on estimated parameters from Monte Carlo simulations on the parameter-derived datasets.

The final estimated parameters were first log-transformed to account for the asymmetric distribution of the utility and probability distortion parameters (ranging from 0 to ∞, with a value of 1 defining the linear case). We then compared the parameter estimates via one-way MANOVA with elicitation condition as main factor. From this multivariate analysis, we identified any significant effect of individual decision functions while recognizing the collective role of all three parameters in capturing risk preferences. More specifically, the MANOVA identified which model function parameters (choice softmax, utility, or probability distortion) differed significantly between CE elicitation conditions.

In the REPEATED condition, the gamble option did not change for long sequences of trials and could, theoretically, be ignored. To test the possibility of an attentional shift toward the safe option in this condition, we defined a model with different weights applied to the two options' values as follows:

$$P_{choose\ Gamble} = 1/(1 + e^{-\lambda((1-k)\times V_G - k\times V_S)}) \qquad (4\text{-}11)$$

The weight parameter (k) captured the attentional shift toward one option, if significantly >0.5. The options' values (V_G, V_S) were computed, as in the previous model, using the power utility function and the selected probability distortion function (Prelec-1).

4.2.5 Evaluation of probability distortion in the Marschak–Machina triangle

We introduced the Marschak–Machina triangle (Machina, 1982; Marschak, 1950) to compare the choice behaviour between the MIXED and REPEATED conditions in an out-of-sample test, and to evaluate the theoretical predictions of the discrete choice model vis-a-vis utility and probability distortions.

The Marschak–Machina triangle defines a 2D space where any probabilistic combination of three fixed reward magnitudes $m_1<m_2<m_3$ can be represented (for details, see Results). The x- and y-axes correspond to the probability of obtaining the lowest (p_1) reward m_1 and the highest (p_3) reward m_3, respectively. The probability of the middle magnitude is not explicitly represented in the diagram, but it can be readily obtained as $p_2 = 1-(p_1+p_3)$. Points on the horizontal axis, therefore, correspond to gambles with outcomes m_1 and m_2, whereas points on the vertical axis identify gambles with m2 and m3 as possible outcomes; the hypotenuse comprises all gambles containing outcomes m1 and m3 only. In our experiment, we set the fixed magnitude levels to $m_1 = 0$ ml, $m_2 = 0.25$ ml and $m_3 = 0.5$ ml.

We characterized monkey A's behaviour within the Marschak-Machina triangle, by defining indifference lines between points on the triangle edges as follows: we presented choices between a fixed gamble (A), defined on one of the axes, and a set of gambles (B_i) located on the triangle's hypotenuse; by fitting a psychometric curve to the ratio of B_i and A choices, we identified the indifference point on the hypotenuse as the probability p_3 corresponding to a choice ratio of 0.5. We then defined an indifference line as the segment connecting the fixed gamble on the axis with its corresponding indifference point. This procedure was repeated for four fixed gambles on the x-axis ($p_1 = 0.2, 0.4, 0.6, 0.8$) and for another four fixed gambles on the y-axis ($p_3 = 0.2, 0.4, 0.6, 0.8$), resulting in 8 indifference lines.

Such indifference lines characterized points on the triangle edges (two-outcome gambles): they did not represent complete indifference curves within the Marschak–Machina triangle (three-outcome gambles). Nevertheless, the slopes of the indifference lines univocally identified a directional property a monkey's risk preference pattern: a gradual change in the slope (fanning-in or fanning-out) of indifference lines has been extensively used in the economic literature to characterize choice behaviour, particularly in relation to the predictions of generalized expected utility theories. This property allowed us to quantify behavioural changes across elicitation conditions and to compare the observed data with predictions from the theoretical economic model.

Crucially, gambles resting on the two axes were never used in the elicitation of CEs, representing an out-of-sample test. As a consequence, the choice behaviour observed in the Marschak–Machina triangle could be used as independent validation for our previous results.

We computed the theoretical indifference lines by calculating, for each of the eight fixed gambles defined above, the probability p_3 for which the theoretical subjective value of the fixed gamble was equal to that of the gamble on the hypotenuse. The subjective value of a two-outcome gamble was defined according to cumulative prospect theory as

$$V(gamble) = u(m_H) \times w(p_H) + u(m_L) \times (1 - w(p_H)) \tag{4-12}$$

where m_H and m_L represent the magnitude of the highest and lowest outcome respectively, p_H the probability of occurrence of the highest outcome, u the power utility function and w the Prelec-1 probability distortion function.

The indifference point was defined as the point on the hypotenuse with subjective value $(u(m_3) \times w(p_3))$ equal to the subjective value of the fixed gamble (V(gamble)). Thus, knowing the value of the fixed gamble, one could identify the indifference point as the probability p_3 satisfying the equation $u(m_3) \times w(p_3) = V(gamble)$ as follows:

$$p_3 = w^{-1}\left(\frac{V(gamble)}{u(m_3)}\right) \tag{4-13}$$

where w^{-1} represents the inverse of the probability distortion function: that is, $w^{-1} = e^{(-(-\ln(w))1/\alpha)}$.

Each daily set of indifference points was elicited after CE elicitation sequences, for both the MIXED and REPEATED CE elicitation sessions. This resulted in two sets of indifference lines, distinctly associated with the REPEATED and MIXED conditions. Both datasets were obtained using intermingled gamble sequences, so any difference in the pattern of indifference lines could only be attributed to the effect of the previous block of trials (i.e., REPEATED or MIXED CE elicitation).

The directional pattern of the indifference lines was characterized by a measure of the 'fanning' direction, corresponding to a gradual change in the slopes of indifference lines. When moving from the lower right to the top left corner of the Marschak–Machina triangle, indifference lines decreasing their slope would fan-in, whereas indifference lines increasing their slope would fan-out, much like the structural slats of a folding fan.

A linear regression analysis on the indifference line slopes was used to statistically characterize the fanning pattern. A positive regression coefficient identified fanning-out of the indifference lines, whereas a negative regression coefficient identified fanning-in. It should be noted that the relation between the slopes of the indifference lines, as we defined them, was not expected to be linear, but the linear regression served as a reasonable description of the expected theoretical pattern and was then used to characterize the measured behaviour.

To statistically compare the predicted and observed sequence effects on the steepness of the indifference lines, we first calculated the shift of indifference points (change in p3 value) between the REPEATED and MIXED conditions; we did this for each of the eight indifference lines, for both the measured data and the model's predicted lines. We then performed a correlation analysis on the modelled and measured shifts.

4.2.6 Trial history effects

Because gamble presentation order was the only difference between the MIXED and REPEATED elicitation sequences, we sought to categorize the effects of said order on the subjective distortion of probabilities. Using past gamble EVs as a quantitative measure of past experiences (specific to probabilities), we compared the distribution and use of previous gamble EVs across elicitation condition.

We first compared the variability of consecutive gamble probabilities in both conditions using a two-sample t-test. We used the absolute value of consecutive gamble EV differences to contrast order in an unsigned manner, as signed differences would amount to 0 in both cases. We then assessed the use of past gamble EVs using the following logistic regression:

$$y = \beta_0 + \beta_1(EV_{Gamble}) + \beta_2(EV_{Safe}) + \beta_3(EV_{Gamble-1}) + \cdots + \beta_n(EV_{Gamble-n}) + \varepsilon \quad (4\text{-}14)$$

Again, the dependent variable took a value of $y = 1$ if the gamble was chosen and $y = 0$ if the safe option was chosen instead. The EV of both the current gamble and safe (EV_{gamble}, EV_{safe}), as well as the gamble EV of up to 8 trials in the past ($EV_{gamble-n}$), served as independent variables. Trials that did not have a minimum of 8 previous trials, in individual sessions, were removed for this analysis. We again standardized regression coefficients and identified how many past gamble EVs had a significant impact on current choice (one-sample t-test, $p<0.05$). Refining the analysis to a singular preceding trial, we investigated the use of a win-stay/lose-shift (WSLS) strategy by the animals. A common strategy for human and nonhuman primates alike, a WSLS choice pattern involves repeating a 'winning' choice until it results in a 'loss,' after which one would shift and try their luck on another choice option. Because choice options in the CE elicitation sequences involved many different values for both the gamble and the safe options, we instead explored a more relaxed WSLS model as follows:

$$y = \beta_0 + \beta_1(EV_{Gamble}) + \beta_2(EV_{Safe}) + \beta_3(\,Outcome_{past}) + \beta_4(Position_{LR}) + \varepsilon \quad (4\text{-}15)$$

If the previous choice had been that of a gamble, and that gamble had won (i.e., resulted in a 0.5 ml reward), the third independent variable ($Outcome_{past}$) took a value of 1; if the past chosen gamble had instead been unsuccessful, $Outcome_{past}$ was 0. By including current EV_{Gamble}, EV_{Safe}, and $Position_{LR}$, we could identify the relative effect of a previous gamble's outcome on current choice. The logistic regression analysis was only applied to trials in which the previous trial's gamble was chosen. A positive regression coefficient for $Outcome_{past}$ implied a greater likelihood of picking the gamble after a 'win', regardless of its value. A negative coefficient would, instead, capture a decrease in the likelihood of picking the gamble after a 'loss.'

To compare the performance of this model with the previously defined model (Eq. 4-2), which did not include the contribution of past trials, we computed the BIC scores of the two models only in trials in which the previous gamble was chosen. After this trial selection, we removed 5 sessions in Monkey A's data, as they had fewer than 4 trials per gamble-safe pair.

To further investigate the effect of past outcomes on the risk patterns, we defined a reinforcement learning model, in which each gamble value was updated, starting from its EV, by adding or removing a fixed amount following a win or a loss, respectively. Formally, choices were evaluated according to the discrete choice model defined earlier (Eq. 4-2), in which the safe value (V_S) was the certain option's magnitude (linear coding of magnitudes), whereas the gamble value (V_G) was updated on each trial according to the following rule:

$$V_G = V_G + \eta \cdot pre_{Win} - \eta \cdot pre_{Loss} \tag{4-16}$$

Where pre_{Win} and pre_{Loss} are variables encoding the last trial's outcome: $pre_{Win} = 1$ if a gamble was won in the previous trial, 0 otherwise, and vice versa for pre_{Loss}. The value-updating parameter η represented the amount of 'value' (in millilitres) added or removed to the value of future gambles based on the prior gamble's outcome. According to this model, the gamble value was not updated if the safe option had been chosen on the previous trial.

We retrieved the η parameter value using MLE, and used the resulting average value to simulate choices and compute the resulting CEs. The simulation was run on MIXED and REPEATED sequences separately, to compare the effect of a value-updating model on the CEs in the two sequence conditions.

4.2.7 Statistical analysis

We used MATLAB and/or Python for all statistical analyses. Logistic regressions were computed per session, and results were standardized by multiplying each coefficient with the ratio of the corresponding independent variable's SD over the SD of the predicted variable (Menard, 2011). Standardized regression coefficients were tested for statistical significance through one-sample t-test. Two-factor ANOVA, one-factor MANOVA, linear regression, and t-test results were considered significant at $p < 0.05$, whereas one-way repeated-measures ANOVAs were Greenhouse–Geisser corrected (degrees of freedom adjustment) to account for sphericity violations (Mauchly's test $p < 0.05$; Greenhouse & Geisser, 1959). Post hoc analysis with Bonferroni–Holm correction for multiple comparisons was applied to ANOVA results. Cohen's d values were used as a measure of effect sizes. In all analyses of data from single sessions, we reported mean ± SEM across sessions.

4.3 Results

4.3.1 Design

We tested whether the shape of the probability distortion would be influenced by the order in which probability information is presented in a sequence of decisions.

Once the animals had been extensively trained with the reward-predicting stimuli (>10,000 trials), we presented them with sequences of binary choices between different probabilistic rewards (or gambles) and safe rewards (Fig. 4-1). We then used the choice ratios to measure the value of gambles relative to certain rewards, pinpointing the certain rewards that were subjectively equivalent to gambles, or a gamble's CE. This procedure revealed the animals' attitude toward risky choices: gamble CEs larger than

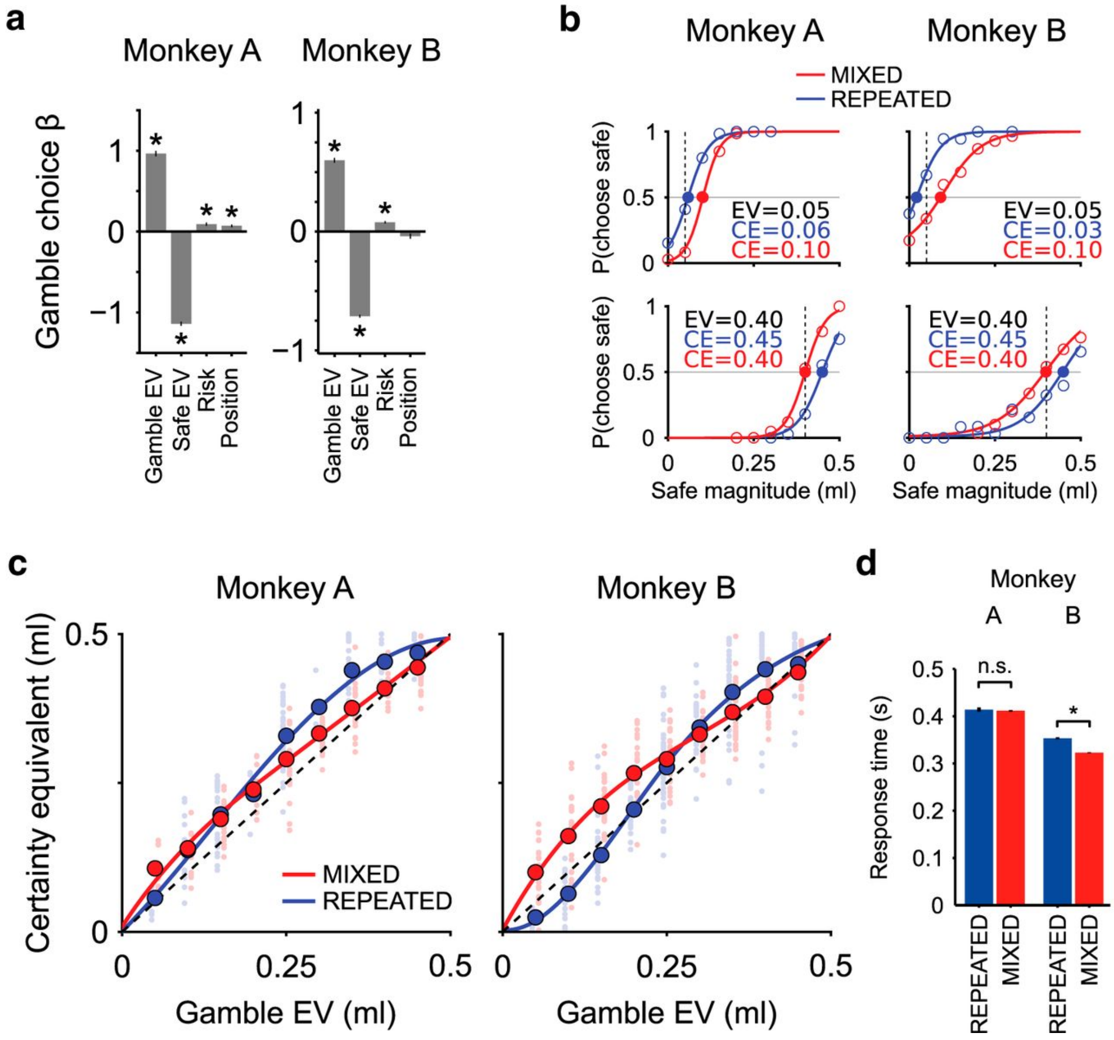


Figure 4-2 | Basic choice behaviour and estimation of CEs. a, Logistic regression of choice behaviour. Four task variables (gamble EV, safe EV [magnitude], risk variance, gamble position) were used as regressors for the gamble choice. Positive standardized coefficients for gamble EV and risk indicated that monkeys preferred gambles with higher EV to gambles with lower EV, and more risky gambles to less risky ones. Negative coefficient for safe EV confirmed that monkeys preferred higher reward magnitudes to lower ones. The positive position factor for 1 monkey indicated a side bias, which was taken into account by repeating all choice pairs with inverted left-right positions.* Significant regression coefficient (one-sample t test). b, Psychometric estimation of CEs. CEs of two example gambles with probabilities 0.1 (top) and 0.8 (bottom), estimated in the two different elicitation sequences: MIXED (red) and REPEATED (blue) sequences. The percentages of safe choices as a function of safe magnitude (circles) were fitted to softmax functions (curves). Vertical lines indicate the gambles EVs (dashed lines). Filled circles represent the CEs. In both monkeys, low probability gambles (top) had a lower CE in the REPEATED condition than in the MIXED condition, where the CEs were consistently higher than the EVs, indicating risk seeking behaviour. High probability gambles (bottom) showed the inverse pattern, indicating risk seeking behaviour only in the REPEATED condition. c, Pattern of CEs across the two elicitation sequences (MIXED vs REPEATED). Single session CEs (small data points) and average CEs across sessions (large data points) plotted as a function of gamble EV, with cubic spline interpolated curves. The full pattern of CEs shows a smooth transition from low to high probability gambles in terms of CE difference across the two elicitation sequences. For low probability gambles, both monkeys showed higher CEs in the MIXED than in the REPEATED conditions; when increasing gamble probabilities, the CE difference across conditions gradually reduced and inverted, resulting in lower CEs in the MIXED than in the REPEATED condition for high reward probabilities. Single session data points were shifted horizontally (REPEATED condition: left; MIXED condition: right) for visualization purposes. d, Mean RT (± SEM across sessions) in the two elicitation conditions. RTs for Monkey A were similar in the two conditions (RT difference = 3.0 ms, t(9088) = −0.59, p = 0.56); Monkey B showed faster response in the MIXED condition compared with the REPEATED condition (RT difference = 30.0 ms, t(22233) = −15.88, p = 1.77 × 10−56) (for RT as a function of the options' EV, see Figure 2-1). *Significant RT difference between conditions (two-sample t test).

said gamble's objective EV reflected risk-seeking behaviour; risk-aversion was characterized instead by gamble CEs smaller than the gamble's EV.

By simultaneously estimating the individual contributions of utility and probability distortion to these measures of risk attitudes, we could model the shape of the monkeys' probability distortion regardless of the utility function.

4.3.2 Basic behavioural performance

A logistic regression analysis demonstrated that the monkeys used the information from the visual stimuli to guide their decisions on all daily testing sessions (Fig. 4-2a). A positive regression coefficient for gamble value (one-sample t-test, Monkey A: $t(55) = 29.41$, $p = 2.5\times10^{-35}$; Monkey B: $t(58) = 30.16$, $p = 3.9\times10^{-37}$) indicated that animals were more likely to choose higher probability gambles than lower probability ones; conversely, the negative coefficient for safe reward value (Monkey A: $t(55) = -44.65$, $p = 6.8\times10^{-45}$; Monkey B: $t(58) = -58.61$, $p = 2.6\times10^{-53}$) indicated that monkeys chose the safe option more frequently when its value was of higher magnitude. Both animals preferred gambles of higher over lower probabilistic variance, i.e., they preferred gambles that were more uncertain, regardless of the outcome (positive coefficient for risk; Monkey A: $t(55) = 4.58$, $p = 2.7\times10^{-5}$; Monkey B: $t(58) = 7.79$, $p = 1.4\times10^{-10}$). Monkey A, but not monkey B, showed a side bias (positive coefficient for the position variable), which was taken into account by balancing the positions of gambles and safe rewards: every option was presented the same number of times on each side of the computer monitor.

4.3.3 Estimation of subjective values using different sequence structures

We used a binary choice paradigm to estimate the monkeys' subjective valuation of specific gambles. We measured the choice ratios between different safe rewards and gambles ranging in probabilities from $p = 0.1$ to $p = 0.9$. Fitting a softmax curve to each of these gamble-safe groups allowed us to estimate the CEs corresponding to different gamble probabilities (see *methods*). These CEs served as a measure of subjective value for unique probabilities and provided us with a precise measure of an animal's risk preference over the range of probabilities tested.

We elicited CEs in both monkeys using two different elicitation conditions: MIXED and REPEATED gamble sequences (Fig. 4-2b). In the MIXED condition, we estimated CEs from sequences of binary choices containing several different gambles pitted against safe rewards. All gamble and safe options presented were randomly intermixed, and multiple CEs were estimated from these sequences, one for each gamble. In the REPEATED condition, we elicited CEs from blocks of trials that contained a single, unique gamble versus different safe rewards. In this way, we elicited a unique gamble's CE for each given block. Importantly, the two conditions used the same visual stimuli; any difference between estimated CEs would therefore be due to the elicitation sequence in which CEs were estimated.

We aggregated the daily CEs of individual monkeys, for both conditions, to determine the risk-preference pattern derived from the CEs measured in each elicitation sequence. The risk- preference pattern

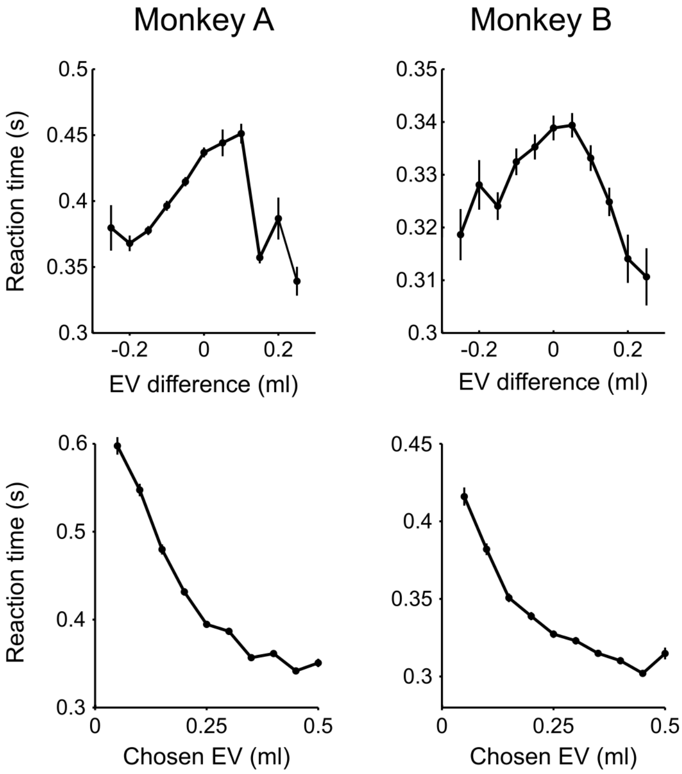


Figure 4-3 | Response time vs EV. Top: Mean RT (± SEM across sessions) as a function of EV difference between the two presented options (gamble EV - safe magnitude). Choices between options with similar EV produced higher RT. Bottom: Mean RT (± SEM across sessions) as a function of the EV of the chosen option. Faster RTs were associated to higher EV of the chosen option, while slower RTs corresponded to choices where a low EV option was selected

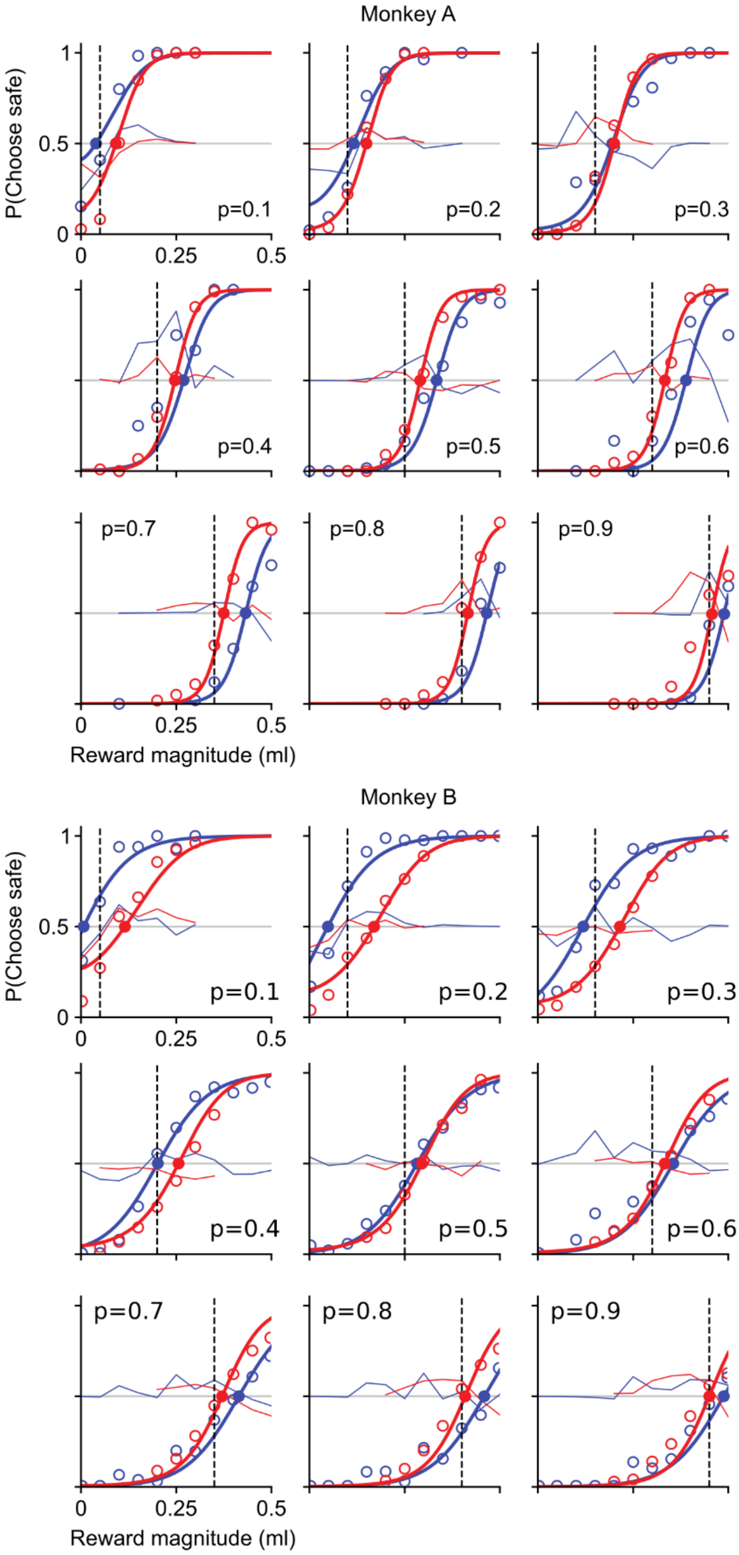


Figure 4-4 | Modelled vs measured choice behaviour. Comparison of estimated (curves) and measured (circles) percentage of safe choices as a function of safe magnitude. Conventions and symbols as in Fig. 4-5d. Thin lines represent differences between estimated and experimental data percentages, with the horizontal line (at 0.5 on the y axis) corresponding to perfect estimate (difference=0).

was therefore directly inferred from the relation between the CEs and the respective EVs, as opposed to being theoretically derived from the shape of utility and probability distortion functions. We found a significant difference between the distribution of CE values elicited in REPEATED versus those elicited in MIXED sequences (two-way ANOVA, factors: gamble probability, elicitation condition). As expected, we found a significant main effect of reward probability on a gamble's CE higher probability gambles had a higher CE in both animals (Monkey A: $F_{(8,237)} = 444.12$, $p = 5.2 \times 10^{-138}$; Monkey B: $F_{(8,337)} = 241.14$, $p = 1.4 \times 10^{-134}$). We also saw a main effect of elicitation conditions (Monkey A: $F_{(1,237)} = 7.69$, $p = 0.006$; Monkey B: $F_{(1,337)} = 20.21$, $p = 9.6 \times 10^{-6}$), where CEs elicited in the MIXED condition were significantly different from those in the REPEATED condition. Adding to this effect, we observed a significant interaction effect between probability and condition (Monkey A: $F_{(8,237)} = 7.73$, $p = 3.3 \times 10^{-9}$; Monkey B: $F_{(8,337)} =$, $p = 8.5 \times 10^{-16}$), suggesting that the different elicitation sequences had a more complex effect on CE values than a mere monotonic increase or decrease. This effect was readily observable from the condition-specific CE distributions (Fig. 4-2c), where the concave pattern of the MIXED-condition CEs contrasts with the S-shaped distribution of the REPEATED-condition CEs. Analysis of the RTs showed no significant difference across conditions for Monkey A, while monkey B responded faster in the MIXED than in the REPEATED condition (Fig. 4-2d). In general, RT observations for both monkeys followed a steady pattern (Fig. 4-3): shorter RTs when choosing higher EV compared to lower EV options, and longer RTs for smaller EV differences between options.

4.3.4 Sequence-dependent changes in probability distortion

Because CE elicitation rested on reward options that varied in both magnitude and probability, any risk-preference changes could be attributed to nonlinear utility, probability distortion, or a combination of both. To better understand the role of these decision variables in shaping a gamble's subjective value, we simultaneously estimated the shape of both functions from the monkeys' daily binary choices. Using a standard discrete choice model (Eq. 4-3), we elicited functional parameters that best explained each animal's choices between gamble-safe choice pairs on individual days, assuming nonlinear utility and probability distortion. The estimation procedure allowed parameters to take on any value, imposing no constraints beyond the functional forms of the discrete choice softmax, probability distortion, and utility curves.

We defined the value of each reward option as the product of its subjective probability and utility, consistent with prospect theory and other modern decision theories (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). As is traditionally done, we modelled utility through a one-parameter power function. The simple power function accounted well for risk-seeking ($\rho > 1$), risk-averse ($\rho < 1$), or risk neutral attitude ($\rho = 1$) for the range of reward magnitudes. We tested only one model for utility, as magnitude presentations did not differ across conditions. Instead, we sought to optimize our choice model with regards to subjective probability because CE elicitation sequences differed in terms of the order in which gamble probabilities were experienced. We tested four classical

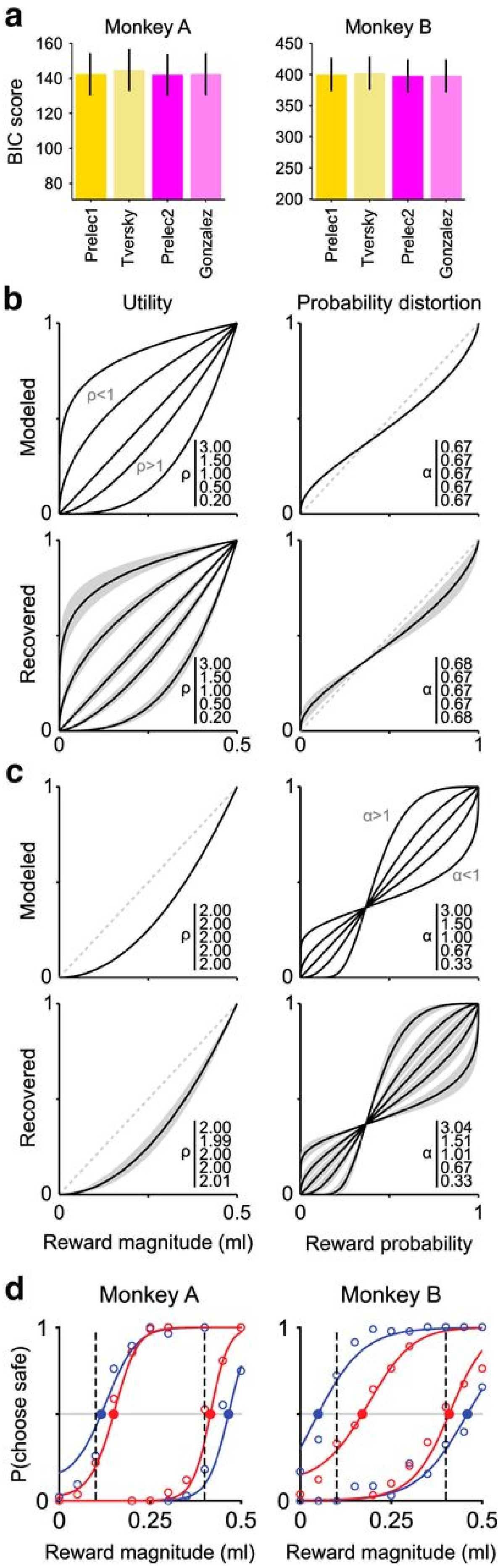


Figure 4-5 | Choice model selection and validation. a, Goodness-of-fit for choice behaviour using four models with different probability weighting functions. Bars represent mean BIC values (±SEM) across all sessions (Monkey A: N = 56; Monkey B: N = 59). BIC scores for daily parametric fits differed significantly across models (one-factor ANOVA with repeated measures, Monkey A: F(3,150) = 8.32, pGGc = 3.1×10^{-3}; Monkey B: F(3,174) = 13.575, pGGc = 5.3×10^{-08}). Lower BIC values for the Prelec weighting functions (Prelec-1, Prelec-2) indicate a better fit of the data compared with the one-parameter Tversky or two-parameter Gonzalez functions. BIC values for all model pairs, except for Prelec-1 versus Prelec-2, Prelec-1 versus Gonzalez, Prelec-2versus Gonzalez in Monkey A, and the Prelec-2 versus Gonzalez in Monkey B, were significantly different (post hoc analysis, $p< 0.05$) for both monkeys. The sum of squared errors in CE estimation was the lowest in the Prelec models. b, c, Validation of the parameter estimation procedure using the Prelec-1 probability weighting function. Top, Utility (left) and probability distortion (right) functions used to simulate choices. Bottom, The functions recovered with the MLE procedure. Monte Carlo simulation of choice behaviour (using the same number of trials and the same step-size for magnitude and probability as in the measured data: 9 gamble probabilities, 11 safe magnitudes, 6 trials per gamble-safe pair) was repeated 1000 times, producing the 95% CIs on the parameter estimates (grey areas). Varying the utility function parameter (ρ, 0.2–3) while keeping the probability distortion parameter constant (α = 0.67) resulted in an unbiased estimate of the utility shape (b). The probability distortion parameter (α), varying from 0.33 to 3 while keeping the utility shape fixed (ρ = 2), was recovered consistently and without bias (c). d, Modelled versus measured choice behaviour. Comparison of estimated (curves) and measured (circles) percentage of safe choices as a function of safe magnitude, for two example gambles (probabilities 0.2 and 0.8) (for the full dataset, see Figure 3-1). Estimated choice percentages were computed using the discrete choice model with the MLE-recovered parameters (Eq. 4-3, using the Prelec-1 probability weighting function). Red and blue points represent estimated CEs. Vertical dashed lines indicate EVs. The estimated psychometric functions closely approximated the measured data points, and differences in estimated CEs across conditions are compatible with the observed data for both low and high probabilities (Fig. 4-2b).

models of probability distortion to maximize the reliability of our model in capturing real choices; two of these functions had one free parameter, and the others had two. Finally, we defined cumulative log-likelihood functions for each of these models and estimated the best-fitting parameters for each decision function through MLE (see *methods*).

Across all testing sessions, the BIC scores of the Prelec curves were consistently lower than the one-parameter Tversky and lower than the Gonzalez models in at least one monkey (Fig. 4-5a). However, while the two-parameter Prelec had a marginally lower BIC score in both animals, the one-parameter Prelec showed marginally lower sums of squared errors between predicted and average experimental CEs (one-factor ANOVA with repeated measures, Monkey A: $F_{(3,144)} = 6.166$, pGGc = 5.7×10^{4}; Monkey B: $F_{(3,168)} = 3.699$, pGGc = 1.3×10^{-2}). We ultimately selected the one-parameter Prelec due to this lower sum of squared errors, lower parameter count, and because of its ease of interpretation: for the curvature parameter $\alpha > 1$, the function underweighed low probabilities and overweighed high ones, for $\alpha < 1$, low probabilities were overweighed and high ones were underweighed. With an $\alpha = 1$, probabilities were treated linearly. Monte Carlo simulations from predefined parameters confirmed the reliability of the MLE method for the selected model: we recovered accurate parameters for both the utility (Fig. 4-5b) and probability distortion (Fig. 4-5c) functions.

Having selected the one-parameter Prelec as the best-fitting probability distortion function, we estimated the functional parameters of our choice model (Eq. 4-3) using the MLE method. The model was able to capture the characteristic pattern of risk attitudes observed in our experimental data: CEs of low probability gambles were larger than their respective EVs in the MIXED condition, whereas CEs of high probability gambles were larger than their EVs in the REPEATED condition (4-5d, see Fig. 4-4 for the full dataset), in accordance with the measured behaviour (Fig. 4-2b).

We compared daily estimated parameters across CE elicitation conditions for utility and probability distortion (Fig. 4-6a). Both animals exhibited convex utility ($\rho > 1$) in the tested range of 0–0.5 ml accounting for risk-seeking behaviour, with linearity only in the case of Monkey B's REPEATED condition. Importantly, probability distortions inverted across elicitation condition. In the MIXED elicitation condition, both animals overweighed low probabilities and underweighed high ones ($\alpha > 1$), whereas they instead underweighed low probabilities and overweighed high ones within the REPEATED condition ($\alpha < 1$) (Fig. 4-6b). MANOVA confirmed the impact of the different elicitation sequences on both animals' choice pattern (Monkey A: $F_{(1,54)} = 24.96$, Wilks's $\lambda = 0.41$, $p = 3.85 \times 10^{-10}$, $\eta^2 = 0.59$; Monkey B: $F_{(1,57)} = 40.78$, Wilk's $\lambda = 0.31$, $p = 5.2 \times 10^{-14}$, $\eta^2 = 0.69$) with only the probability distortion parameter (α) consistently different across conditions (Fig. 4-6a,c). The change in risk-attitude between the two conditions could therefore, at least in the case of gamble-safe choices, be reduced to a reversal in the probability distortion function.

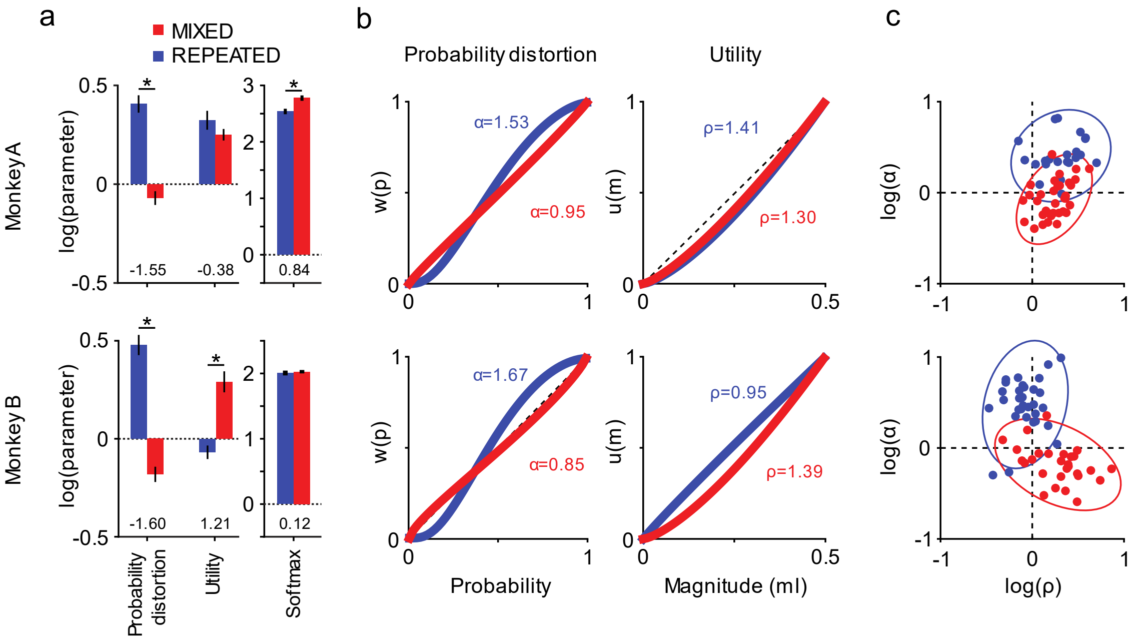


Figure 4-6 | Utility and probability distortion functions in two elicitation conditions. a, Model parameter estimates (mean ± SEM across sessions) in the MIXED (red) and REPEATED (blue) conditions. *Significant differences across conditions (MANOVA). The probability distortion parameter (α) consistently varied across sequence structures in both monkeys: negative log values in the MIXED condition corresponded to inverse S-shaped probability distortion ($\alpha < 1$), whereas positive log values in the REPEATED condition implied S-shaped probability distortion ($\alpha > 1$). Numbers below the bars indicate effect sizes (Cohen's d). The utility (ρ) and softmax (λ) parameters significantly differed across conditions only for 1 monkey, with a smaller effect size compared with the probability distortion parameter. b, Shapes of the probability distortion function (left) and utility function (right) corresponding to the estimated parameters, displaying the consistent difference in subjective probability evaluation across conditions for both monkeys. c, 2D representation of the utility and probability distortion parameter estimates. Dots represent the simultaneously estimated utility (ρ) and probability distortion (α) parameters for single behavioural sessions, with 95% confidence ellipses.

The REPEATED condition was a much less complex decision situation compared with the MIXED one, theoretically allowing for a simpler choice strategy: it would have been sufficient to evaluate the certain option, ignoring the gamble option in the majority of trials, to make choices. We tested for this possibility by fitting a model with an attentional parameter to the choice data (Eq. 4-11). We found that there was no significant difference in attention given to the safe compared with the gamble option (the weight parameter was not significantly different from 0.5; Monkey A: $t_{(21)} = -2.01$, $p = 5.7 \times 10^{-2}$ (t-test); Monkey B: $t_{(30)} = -1.25$; $p = 2.2 \times 10^{-1}$), suggesting that both options were fully considered when making choices in the REPEATED condition. Furthermore, shorter RTs in the REPEATED condition, expected if the monkeys ignored the gamble option, were not observed (Fig. 4-2d).

4.3.5 Reversal of probability distortion in the Marschak–Machina triangle

To extend our findings beyond gamble-safe choices, we characterized the choice behaviour of 1 monkey in a different set of gambles using the Marschak–Machina triangle. This diagram was first introduced as a way of 'organizing' a series of anomalies observed in risky choices, most notably the common ratio and common consequence effects, which violated the independence axiom of EU theory (Allais, 1953). Several economic theories were developed to explain these apparent paradoxes. Each theory predicted indifference curves with distinctive shapes in the Marschak–Machina triangle, making it an ideal framework to evaluate and compare the alternative theories (Machina, 1982).

The use of this diagram, which makes it possible to represent a more general class of choice options (i.e., gambles with three fixed outcomes of varying probabilities; Fig. 4-7a), allowed us to extend our results to a wider range of problems. We did this to test the robustness of the parametric modelling (out-of-sample test) and, most importantly, to investigate the effect of elicitation condition from a different perspective: by looking at the change in direction of indifference lines, which connected points of the triangle edges (specific two-outcome gambles) for which the animal expressed choice indifference (Fig. 4-7b), we could quantify the effects of elicitation condition that were specifically dependent on changes in probability distortion, and independent of changes in the shape of the utility function.

One of the theoretical consequences of probability distortions in the Marschak–Machina triangle is that indifference lines would not be parallel to each other, as in the case of linear probability weighting, but would instead fan-out or fan-in depending on the probability distortion (Fig. 4-7c): an inverse S-shaped probability distortion would induce fanning-out, whereas an S-shaped one would result in indifference lines fanning-in. Fanning-out would indeed correspond to an increase in the steepness of the indifference lines when shifting 'probability mass' from worse to better outcomes. As steeper lines correlate with more risk-seeking behaviour, fanning-out would imply an inverse S-shaped probability distortion. The opposite would happen with fanning-in indifference lines, then corresponding to an S-shaped probability distortion function (Camerer, 1989). Crucially, because the outcome magnitudes used in the Marschak–Machina triangle are fixed, the fanning direction is independent of the utility function and is therefore solely determined by the shape of the probability distortion. In that sense, any observed change in the fanning direction of the indifference lines with a change in elicitation sequence could only be due to a change in the probability weighting function (Fig. 4-7c).

We used the previously recovered parameters for utility and probability distortion to estimate the expected pattern of indifference lines in the two experimental conditions: MIXED and REPEATED sequences. We then compared the predicted directions of the indifference lines with the measured ones. As expected, the theoretical indifference lines, modelled using the previously elicited parameters, showed a slight fanning-out pattern for the MIXED condition, where a weakly inverse S-shaped probability distortion was measured. Conversely, we saw a fanning-in pattern in the REPEATED condition, for which we had observed an S-shaped probability distortion (Fig. 4-8a, left).

The direct experimental measure of indifference lines was performed by presenting the animals with binary choices between a gamble represented by a fixed point on the triangle edge and one of several points on the triangle's hypotenuse. The indifference line was defined as the segment connecting the fixed point with the point corresponding to choice indifference on the hypotenuse. This procedure resulted in a directional pattern of indifference lines compatible with the theoretically predicted one, with no clear fanning direction of indifference lines in the MIXED condition, and clear fanning-in in the REPEATED condition (Fig. 4-8a, right). We quantified this directional pattern of indifference lines using a measure for the fanning direction. The fanning of indifference lines corresponds to a gradual change in the slope of indifference lines: when moving from the lower right corner of the probability triangle to the upper left corner, an increasing slope would produce fanning-out, whereas a decreasing slope would produce fanning-in. Following this principle, we statistically assessed the fanning direction of the indifference lines by computing a linear regression on the slopes of the indifference lines. Results show no significant regression slope in the MIXED condition ($R^2 = 0.08$, $p = 0.50$), suggesting no fanning of indifference curves, whereas in the REPEATED condition a significant linear regression ($R^2 = 0.98$, $p = 4.4 \times 10^{-6}$) indicated fanning-out of the indifference lines. These results are consistent with predictions from the modelled indifference lines, which show a similar pattern of fanning directions (Fig. 4-8b).

We statistically compared the measured and predicted patterns of indifference lines by calculating the shift in the location of indifference points across conditions, corresponding to changes in the slope of indifference lines. A significant correlation between the predicted and measured shifts (Pearson's correlation coefficient $r = 0.78$, $p = 4.0 \times 10^{-3}$) confirmed that the experimental data complied with our theoretical predictions (Fig. 4-8c) and supported the finding that probability distortion drove the change in risk attitude between REPEATED and MIXED conditions.

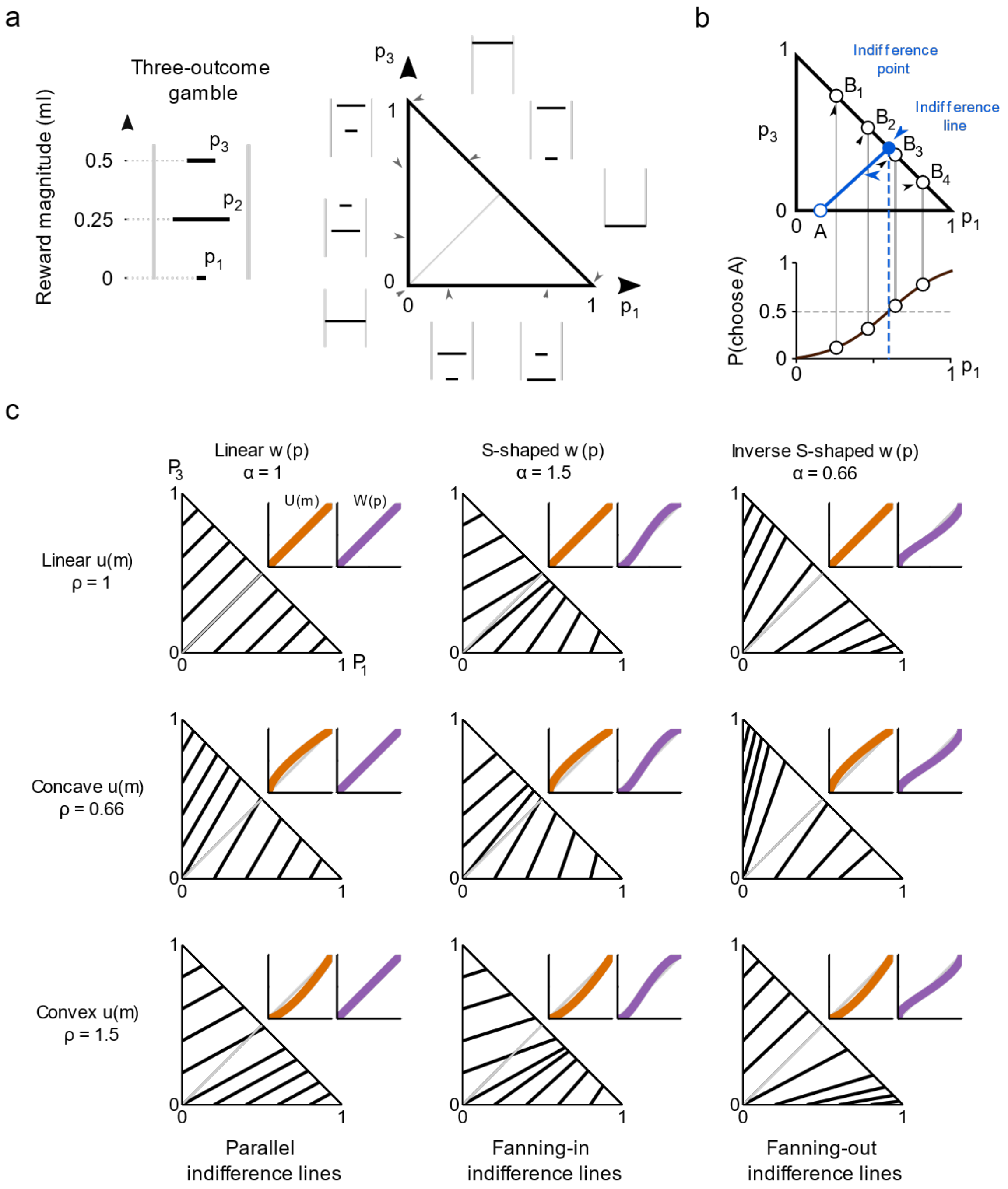


Figure 4-7 | Indifference lines in the Marschak–Machina triangle modeling different patterns of probability distortion. a, Representation of gambles in the Marschak–Machina triangle. Schematic representation of a three-outcome gamble (left): probabilistic combination (p1, p2, p3) of three fixed magnitudes (m1 = 0 ml, m2 = 0.25 ml, m3 = 0.50 ml), which can be represented in the Marschak–Machina triangle (right, with example gambles corresponding to points on the triangle edges). Grey line in triangle connects points with equal EV (0.25 ml). b, Procedure for the psychometric measurement of one indifference line. An indifference point (top, blue dot) in choices between a fixed gamble A and different gambles Bi (circles) was defined as the point on the triangle hypotenuse for which a softmax function fitted on the ratio of A over Bi choices equated 0.5 (bottom). An indifference line was then constructed by connecting such indifference point on the hypotenuse to the fixed gamble A (blue line). c, Theoretical indifference lines. Indifference lines predicted by cumulative prospect theory, for different underlying shapes of utility (u(m), power function) and probability distortion (w(p), Prelec-1 function). Each plot represents the indifference lines corresponding to a particular combination of u and w shapes, represented by orange and purple lines, respectively. The shape of the utility function (linear in the first row of plots, concave and convex in the other two rows) changes the global orientation of the indifference lines, without affecting their fanning direction. On the contrary, a change in probability distortion shape corresponds to a change in the fanning direction of indifference lines: a linear probability distortion (first column) produces parallel indifference lines, whereas S-shaped (second column) and inverse S-shaped (third column) probability distortions correspond to indifference lines fanning-in and fanning-out, respectively, regardless of the utility function shape.

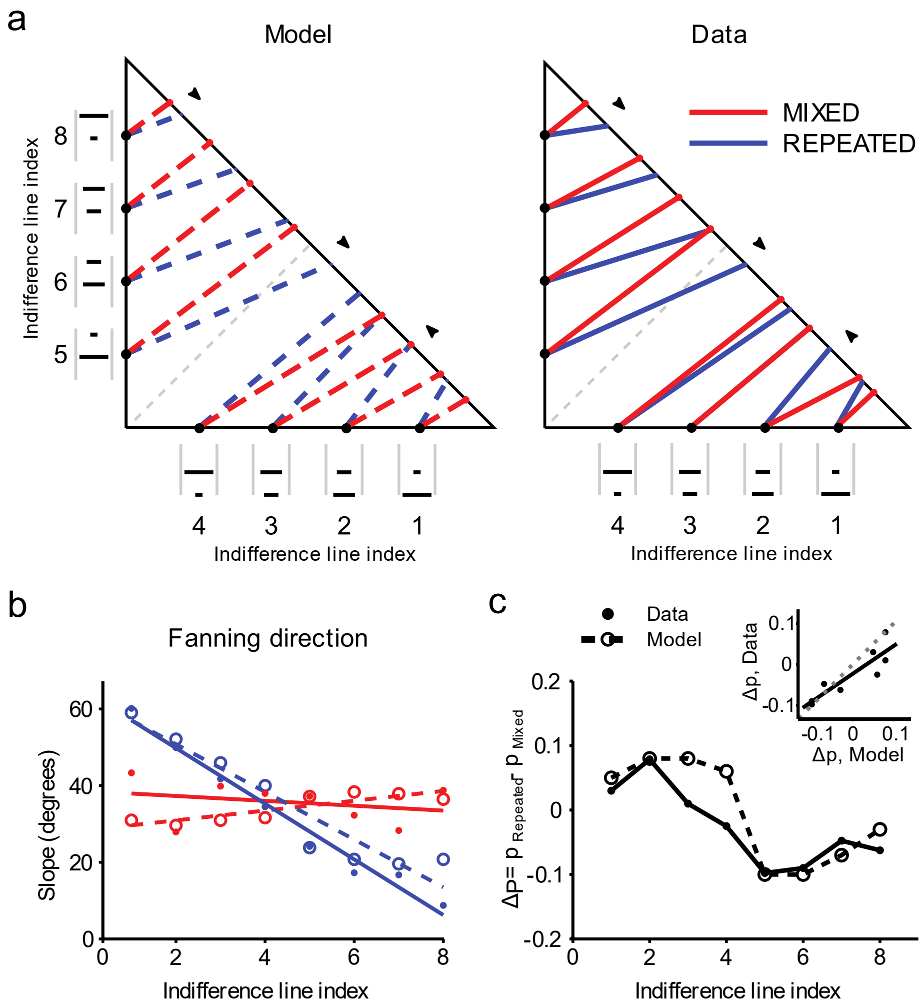


Figure 4-8 | Effect of CE elicitation sequences on the Marschak–Machina triangle indifference lines. a, Modelled (left) and measured (right) patterns of indifference lines across conditions. Arrows indicate the direction and amount of shift for three sample indifference points between the MIXED (red) and REPEATED (blue) conditions, highlighting how the model correctly predicted the effect of condition change. Grey line connects points with the same EV (0.25 ml), representing an indifference line in case of risk-neutral behaviour. Numbers define indices for the indifference lines, corresponding to fixed gambles on the triangle edges (black dots, also represented as visual cues). b, Fanning direction of the indifference lines. Points represent the slope of indifference lines (angle between each line and the horizontal axis) as a function of indifference line index. Circles represent the model predicted values. Dots represent experimental data. Lines indicate linear regressions, separately computed on the two task conditions for the model (dashed lines) and the data points (continuous lines). A regression line with negative slope corresponds to a decrease in indifference line angle, indicating fanning-out; conversely, a positive regression coefficient indicates fanning-in of indifference lines. c, Statistical comparison between model and experimental data. Shift in location of indifference points across elicitation sequences (average difference ± SEM). A linear regression between the modelled and measured shifts (inset) confirmed the match between model and data in terms of predicted shift in indifference points, corresponding to a correct prediction of the change in the fanning direction across conditions.

4.3.6 The effect of trial history on the probability distortion

Because the structure of elicitation sequences appeared to affect probability distortions specifically, we investigated whether the differences in choice behaviour could be explained in relation to past experiences, or trial history. One key difference between elicitation sequences was the order of the probabilities presented on the screen. In the MIXED sequences, the monkeys were much more likely to have experienced different gambles in their immediate past than in trials from REPEATED sequences, where the same gamble was repeated numerous times. Consequently, while the range of probabilities, magnitudes, and safe outcomes was identical in both conditions, the variability of past gambles was significantly different between the two conditions (Fig. 4-1d,e).

Because human and nonhuman primates, much like rodents, often base part of their risky decisions on recent experiences (G. Barron & Erev, 2003; Hayden et al., 2008; Marshall & Kirkpatrick, 2013; Nowak & Sigmund, 1993), we again ran a logistic regression on the probability of choosing the gamble option: this time to verify whether the EV of past gambles had any impact on the animals' decisions (Eq. 4-14). We found that, in the MIXED condition, both monkeys made use of at least one past gamble to make their decision (Fig. 4-9a). The monkeys appeared to bias their choices in favour of the gamble (positive regression coefficient) when the prior gamble's EV was higher. In game-theoretic terms, and taking the gamble's EV as a proxy for its ‘win rate,’ monkeys seemed to follow a WSLS strategy, whereby receiving a reward from a risky choice option increased the likelihood of choosing a similar option again; the opposite was true for choices where the risky option resulted in a loss (no reward). To validate this hypothesis, we applied a WSLS-compatible model (Eq. 4-15) on the immediate trial history of both monkeys, looking at their propensity to choose a gamble over a safe outcome when they had previously chosen a gamble and won (Fig. 4-9b). As expected, we found a significant effect of both the current gamble's EV (one-sample t-test, Monkey A: $t_{(50)} = 29.41$, $p = 3.19 \times 10^{-33}$; Monkey B: $t_{(58)} = 32.28$, $p = 9.38 \times 10^{-39}$) and the current safe outcome's EV on the likelihood of choosing a gamble (one-sample t-test, Monkey A: $t_{(50)} = -38.71$, $p = 6.05 \times 10^{-39}$; Monkey B: $t_{(58)} = -46.19$, $p = 1.9 \times 10^{-47}$). Both monkeys had a small but significant side bias (one-sample t-test, Monkey A: $t_{(50)} = -4.59$, $p = 2.97 \times 10^{-5}$; Monkey B: $t_{(58)} = -2.55$, $p = 1.3 \times 10^{-2}$). More importantly, there was a significant positive effect of ‘winning’ the preceding gamble on the likelihood of selecting the gamble option again, regardless of value (one-sample t-test, Monkey A: $t_{(50)} = 10.75$, $p = 1.3 \times 10^{-14}$; Monkey B: $t_{(58)} = 8.32$, $p = 1.76 \times 10^{-11}$). In other words, receiving a reward from a risky gamble made the next gamble more attractive relative to the safe outcome.

We investigated this effect further, by estimating separate utility and probability distortion parameters in trials where a past gamble had been selected and ‘won’ and in trials where the past selected gamble had been ‘lost.’ Due to lower trial counts per session after this trial selection, all sessions were pooled for each condition. In both animals, the utility function estimated from the former class of trials was more convex than the utility estimated from unrewarded trials (Fig. 4-9c). Probability distortions,

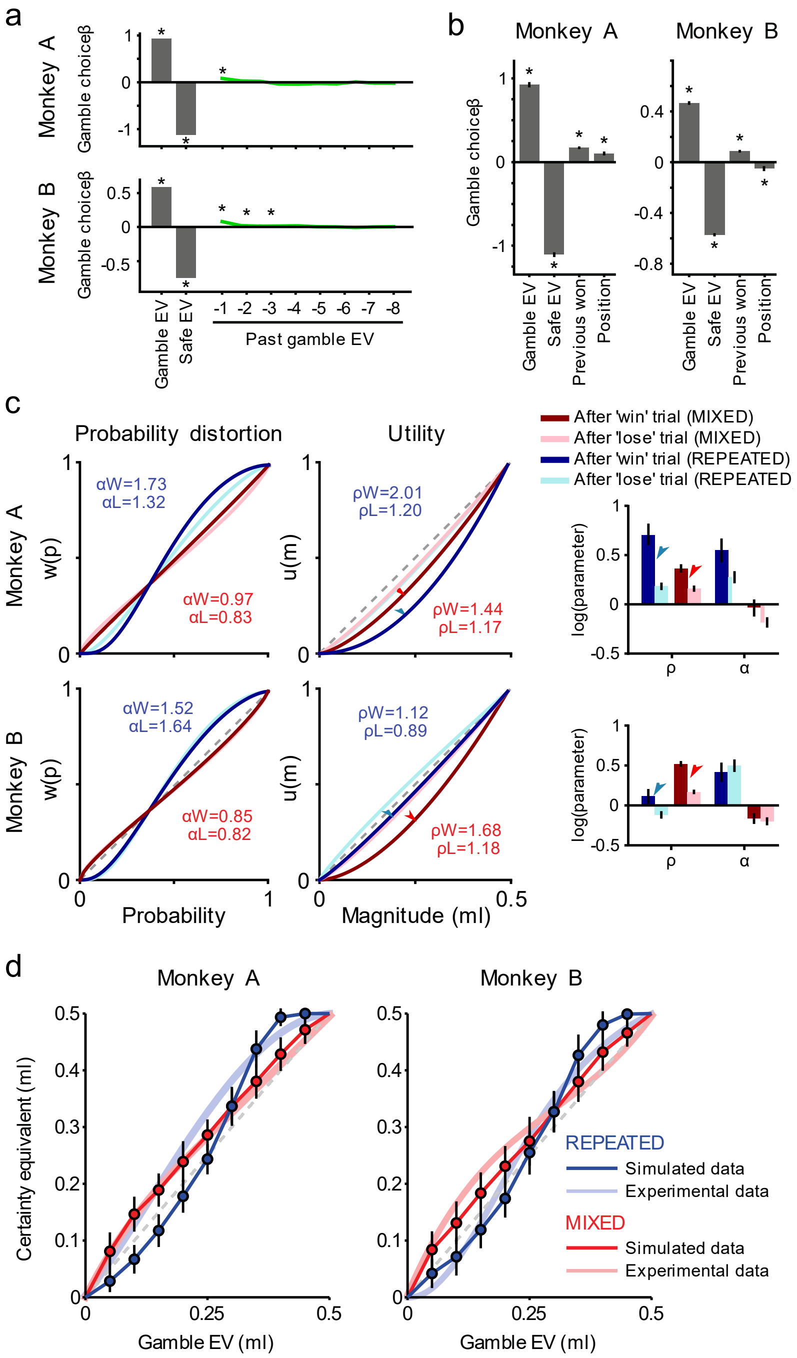


Figure 4-9 | Sequence-dependent effects of trial history on probability distortion shape. a, Influence of past trials on current trial's choice. Standardized regression coefficients (mean ± SEM across sessions) for current trial's gamble EV, safe reward magnitude, and previous trials' gamble EV (up to eight trials in the past). *Coefficients significantly different from zero (t test). For both monkeys, the choice behaviour depended on at least one trial in the past. Positive regression coefficients indicated that an increase in the previous trial's gamble EV induced the monkeys to choose the current trial's gamble option more frequently. b, Effect of the past outcomes on gamble choices. Standardized regression coefficients (mean ± SEM across sessions) for gamble EV, safe magnitude, previous trial's gamble outcome (0 or 0.5 ml), and gamble position. A significant positive coefficient for the previous outcome indicated that monkeys chose the gamble more often when the previously chosen gamble was successful (0.5 ml) than when it was not successful (0 ml): the gamble was chosen more after a win than after a loss. In terms of BIC score, this model (Eq. 4-15) was at least as good at describing the choice data compared with the model with no past trials' influence (Eq. 4-2) (Monkey A: BIC2 = 84.2, BIC14 = 82.3, t test: p = 0.14; Monkey B: BIC2 = 221.4, BIC14 = 215.8, t test: p = 5.8 × 10−5). c, Effect of past outcomes on the utility and probability distortion functions. The utility function appeared more convex following a gamble-win trial (0.5 ml reward) than following a loss (no reward), suggesting that gamble outcomes had an influence on the relative value of gamble and safe options on the next trial. The utility parameter estimates following win and loss trials are indicated as αW and αL, respectively, whereas probability distortion parameter as ρW and ρL, respectively. Arrows indicate the change in the utility parameter between loss and win trials. Error bars indicate the 95% CIs of the parameter estimates. d, Simulated effect of EV update mechanism based on past outcomes. Mean ± SEM across simulated sessions (N = 50) of the CE resulting from choices simulated using the learning model (Eq. 4-16) in MIXED and REPEATED conditions. The parameters used in the simulation were recovered from the MLE procedure with the same model separately for each monkey. Linear probability weighting and linear magnitude coding were used in the simulation, demonstrating that an EV update mechanism interacting with the local trial structure could explain the observed change in risk attitudes across conditions without explicitly introducing a nonlinear probability distortion.

however, were not consistently different between these two classes of trials, maintaining their respective inverse-S and S shapes for MIXED and REPEATED conditions. Much like in the logistic regression, these results suggested a tendency to choose the gamble option more often after rewarded (win) trials, compared with unrewarded trials (a more convex utility function corresponding to stronger risk-seeking behaviour). What it also highlighted, however, was a change in the relative value distribution between gambles and safe options: one that varies with past experience. In other words, gambles following a rewarded trial would be of higher relative value for the monkeys than those following unrewarded trials, at least in terms of safe rewards.

The effect of past wins or losses on subjective value could account for some of the gap in probability distortion observed across our two conditions. A MIXED sequence of gambles would drive subjective value estimates in an opposing pattern to that of a REPEATED elicitation sequence simply due to task structure. In the case of MIXED sequences, the random distribution of gamble probabilities would indeed result in an inverse-S probability distortion. Gambles with probabilities >0.5 would, more often than not, follow a gamble of lower EV; the monkey would then, on average, be less likely to pick said gamble due to the decrease in subjective value estimate following lower past returns. This would drive down the CE value of high probability gambles. In the case of low probability gambles, high past returns would drive CEs up. From this, we would expect an opposing distortion pattern in a REPEATED condition. For any gamble, the CE value would be distorted in a way proportional to its own probability: a low probability gamble would be driven down in value by repeated experience, whereas a high probability gamble would see its value go up. A change in gamble value, rather than a simple WSLS strategy, might also have longer lasting effects and could explain the persistence of sequence-type effects when looking at choices in the Marschak–Machina triangle paradigm.

To test this hypothesis directly, we developed a simple reinforcement learning model in which gamble values were updated based on the previous trial's outcome: the value of a gamble increased by a fixed amount after a win and decreased by the same amount after a loss (Eq. 4-16). Importantly, in the choice model, the gambles' starting values were the respective objective EVs, which were compared with the objective safe magnitudes to make choices. No utility or probability distortion was included, only the previous choice softmax function, and we made no distinction between parameters estimated in repeated or mixed sequences. We again estimated the model parameters through MLE on each session's trial-by-trial choice data and retrieved a significant, mean value-updating parameter for both monkeys (Monkey A: $\eta = 4.5 \times 10^{-3} \pm 9.0 \times 10^{-4}$ SEM; $t_{(55)} = 4.96$, $p = 7.1 \times 10^{-6}$; Monkey B: $\eta = 4.1 \times 10^{-3} \pm 5.8 \times 10^{-4}$ SEM; $t_{(58)} = 7.1$, $p = 2.0 \times 10^{-9}$). The value of η corresponded to the fixed amount of value being added to or removed from the gamble's subjective value estimate following ‘win’ and ‘lose’ trials, respectively.

After running the estimation procedure on all sessions, we tested whether the average observed value-updating parameter could explain the different CE distributions seen across the MIXED and REPEATED conditions. We computed CEs from simulated choices using the learning model defined above (Eq. 4-16), using the mean recovered softmax and value-updating parameters, still holding utility and probability weights linear. The resulting pattern of simulated CEs (Fig. 4-9d) followed the experimental pattern. In particular, it captured the clear separation between the two CE elicitation sequences. Although this model appeared to have a higher BIC score than the ‘classical’ prospect theory model (Eq. 4-3; Monkey A: $BIC_{Eq16} = 160.7$, $BIC_{Eq3} = 137.5$, $t_{(55)} = 6.92$, $p = 5.01 \times 10^{-9}$; Monkey B: $BIC_{Eq16} = 419.8$, $BIC_{Eq3} = 392.7$, $t_{(58)} = 4.69$, $p = 1.70 \times 10^{-5}$), it accounted for the change in the pattern of CEs across both conditions using a single set of parameters. Conversely, two different sets of parameters were necessary for the prospect theory counterpart to capture the different CE patterns.

Together, these results suggest that a simple value-updating mechanism that modifies gamble values based on the previous outcomes, applied to different elicitation sequences, would be sufficient to induce a reversal in the observed probability distortion patterns of monkeys during choice.

4.4 Discussion

This study demonstrated that the shape of the probability weighting function guiding value-based choices in monkeys depended largely on the task's sequence structure. When deriving CEs from sequences in which different probabilistic rewards pseudorandomly alternated (MIXED), we found that monkeys overweighed low probability rewards and underweighed high probability ones. Conversely, the same CE elicitation method yielded the opposite choice pattern (underweighting of low probabilities and overweighting of high ones) when choice sequences consisted of trial blocks each containing a unique, REPEATED gamble. By simultaneously eliciting utility and probability weighting functions from each of these elicitation conditions, we showed that the two opposing choice patterns we observed could be explained by a reversal of the standard inverse S-shaped probability distortion function, seen when gambles were MIXED, to an S-shaped distortion when identical gambles were REPEATED. We confirmed and extended these results by comparing choice indifference lines in the Marschak–Machina triangle representations of the two elicitation conditions. The triangle's indifference maps were compatible with the observed inversion of probability distortions, preserving the weighting patterns in trials where no safe options were presented. Finally, by analyzing both sequence structure and monkeys' choices in relation to previous trials, we showed that a past-driven update of subjective values could partially explain the observed reversal in probability distortion.

Modern economic theories of choice under risk introduced distorted probability weightings to account for biases and departures from expected utility theory's predictions (Allais, 1953; Kahneman & Tversky, 1979; Von Neumann & Morgenstern, 1944). Since then, the typical finding has been that humans overweighed low probabilities, all the while underweighting high ones (Abdellaoui, 2000; Gonzalez & Wu, 1999; Lattimore et al., 1992; Tobler et al., 2008): an inverse-S probability distortion (Kahneman & Tversky, 1979). This shape has also been replicated in monkeys (Stauffer et al., 2015) where human-ported tasks resulted in a reliable inverse-S probability distortion. The current study ties in with these findings, using a coherent set of visual stimuli for both gambles and safe reward options to control for any bias introduced by the different visual representations of the two option types. Our results, in addition to reliability capturing macaque behaviour using modern economic choice theories, further characterize the effects of sequence structure on utility and probability distortion.

In contrast to the generally reported inverse-S-shaped probability distortion, a growing number of studies on human and animal subjects have highlighted the variability in probability distortion shapes, both across subjects and between task conditions (Bruhin et al., 2010; Farashahi et al., 2018; Hey & Strazzera, 1989). Recent work by Farashahi et al. (2018) emphasized the flexibility of probability weights in adapting to contextual changes, after finding that S-shaped and linear probability distortions could be elicited in monkeys when performing different tasks. Our experimental data confirmed this high level of behavioural flexibility in monkeys, whereby directly manipulating the order of presented gambles in a single task produced opposing patterns of probability distortion.

Other findings from human experiments suggest that the way in which probability information is presented could account for the reported variability in subjects' risk attitudes. For example, when reward probabilities are explicitly described (choice from description) to human subjects, they act as if overweighting the probability of rare events, but when probabilities are learned from experience (choice from experience), subjects choose as if underweighting the probability of rare events. This effect has been aptly referred to as the description-experience (DE) gap (Hertwig et al., 2004) and appears to extend to other primates. Indeed, monkeys have been shown to be more risk-seeking for experienced than for described gambles, suggesting a similar DE gap effect in nonhuman primates (Heilbronner & Hayden, 2016). Whereas some authors have called for two separate theories explaining choices from description and choices from experience (Abdellaoui et al., 2011; Hertwig & Erev, 2009), others have suggested that prospect theory could effectively describe choice in the two situations when allowing for a change in the probability distortion function between the two settings (Frey et al., 2015; Ungemach et al., 2009).

While the dichotomous choice patterns we observed are comparable with those described in the DE gap studies, here the cues representing reward probabilities were identical in the two sequence conditions. In both MIXED and REPEATED sequences, probabilities were described explicitly through cues, learned from experience by the animals; the conditions only differed in the presentation order of the probability information. While the task design was different from previous human DE studies in this respect, the repeated sampling of outcomes typically used to ‘learn’ the value of risky prospects in choices from experience (for review see Wulff, Mergenthaler-Canseco, & Hertwig, 2018) echoes the repetitive structure of our REPEATED sequence; conversely, described prospects are typically presented in a less structured, randomized sequence, analogous to our MIXED condition. While a direct comparison remains to be made, findings in both the DE gap experiments and in the present study suggest that past trial outcomes play a role in shaping the subjective perception of reward probabilities.

Sampling bias has been identified as a source of variability in probability distortions, particularly in relation to the DE gap. Indeed, sampling bias is particularly problematic in ‘experienced’ conditions due to the limited number of trials used in learning the options' values: with small sample sizes, low probability gambles are often rewarded less frequently than would be prescribed by their nominal probability (Camilleri & Newell, 2013; Hertwig & Erev, 2009; Hertwig & Pleskac, 2010). The use of identical descriptive cues and elicitation procedures in the present study ensured that similar sampling sizes were applied, and indeed required, to estimate CEs for every gamble. Any bias would therefore affect the two conditions in a similar manner. With no obvious sampling biases, our data suggest that the DE gap could be modelled on the probability distortion changes we observed across task conditions, and that much like in the present study, the observed changes in risk-preferences from described to experienced reward probabilities, might result from differences in the task's presentation order of probability information.

A final source of variability we considered was that the REPEATED condition was a much less complex decision situation than the MIXED one: one could ignore the gamble in long, repeated sequences. However, we found that the animals neither differentially weighed the options nor made choices faster in the REPEATED condition, indicating that they were not using widely differing valuation strategies.

The Marschak–Machina triangle, a diagram widely used in the economics literature, allows for the intuitive representation of choices between two- and three-outcome gambles, serving as an ideal framework for investigating complex economic choice problems (Camerer, 1989; Machina, 1987). In the current experiment, we elicited indifference points in the Marschak–Machina triangle representation of the monkeys' behaviour, which crucially provided a link between animal and human studies. Although full indifference curves within the Marschak–Machina triangle remain to be tested, we showed that indifference points on the triangle edges complied with economic theories of choice, and confirmed the reversal of probability distortion across conditions, this time with probabilistic rewards only. Consequently, we demonstrated the possibility of rigorous behavioural characterization in nonhuman primates, paving the way for future investigations into the neurophysiological basis of advanced economic constructs, such as probability distortion, specific economic axioms, or the neural counterparts of alternative economic theories.

In conclusion, our results demonstrated the effect of a task's sequence structure on the shape of a monkey's elicited probability distortion, and highlighted the potential influence of past rewards on subjective value. Moreover, and perhaps most significantly, these adaptive effects extended through time: the patterns of indifference lines observed in the Marschak–Machina triangle after a session of MIXED or REPEATED sequences were compatible with the probability distortion shapes measured in the preceding CE elicitation session, even though the paradigm used in the Marschak–Machina triangle was always randomized. In this sense, the CE elicitation sequences preceding the Marschak–Machina triangle paradigm might have driven and reinforced a gap between the subjective values of identical probabilities, one that influenced choices between gambles in the Marschak–Machina triangle. The reinforcement learning model we used supports this hypothesis, implying that each experienced outcome could reinforce and update the subjective value of probabilities, leading to a flexible and contextually driven judgment of probabilistic information. More sophisticated models, such as the addition of a standard Rescorla–Wagner learning rule or a nonlinear transformation of safe magnitudes to the current value updating mechanism, could be more biologically plausible and successful in explaining the choice mechanism, and remain to be explored. It should be noted that the monkeys' initial learning/association phase was not analyzed here in reinforcement learning terms, as it was performed with imperative trials. A better understanding of probability learning, and the permanence of subjective values reinforced across different conditions, could shed light on the core elements of prospect theory and on the undeniably adaptive nature of utility and probability distortions.

Chapter 5

General discussion

5.1 Summary of findings

The body of work presented in this thesis sought to answer one simple question: what if preferences were adaptive? From that original question, what emerges is that far from having the stable and fixed preferences prescribed by traditional economic models, rhesus macaques flexibly adapt their preferences in a way that optimizes their decisions given the demands of the task at hand.

Specifically, the results in chapters 2-4 highlight the role of contextual adaptation in shaping monkeys' attitude towards risk. In Chapter 2, I demonstrated that the range of rewards that monkeys had experienced in the recent past significantly influenced their propensity to take risks. When the monkeys were presented with new reward magnitudes that they had never experienced, their preferences shifted almost instantly to account for the new, expanded realm of possibilities. In Chapter 3, I uncovered important differences in the way monkeys evaluated the magnitudes of risky (uncertain) and riskless (certain) rewards – even after accounting for the effects of Prospect Theory's probability distortion. In fact, more than just a risky/riskless dichotomy, the monkey's choice preferences in one risky task did not match their choice preferences in another risky task. Finally, where chapters 2 and 3 described the factors that led preferences to adapt, the results of Chapter 4 demonstrated that using a single adaptive model could reconcile what economics would have traditionally recognized as two static ones. By changing the order in which rhesus macaques experienced probabilistic gambles, I demonstrated that adaptation to recent experiences was enough to induce a full reversal of a monkey's preference for risk (at least in the way economics measures preferences). In doing so, I also provided an alternative, adaptive explanation for one of the most puzzling findings in behavioural economics: the description-experience gap[30] (Hertwig & Erev, 2009; but see Wulff, Mergenthaler-Canseco, & Hertwig, 2018).

5.2 Neuroethology at the centre of a paradigm shift

Many attempts have previously been made, particularly by neuroeconomists, to reinterpret top-down decision models as a direct representation of brain activity (Glimcher & Fehr, 2014; O'Doherty, 2014). In doing so, however, most have ignored a crucial question: do these interpretations also make sense

---

[30] The fact that people's attitude towards probabilistic rewards reverses depending on if probabilities were stated or described, versus if they had been learned from experience

from a bottom-up, evolutionary perspective? In other words, how does economics tie-in with neuroethology?

As ironic as it seems, biologists have often been the main culprits in propagating unrealistic top-down interpretations of economic findings; doing so when they are, in fact, uniquely positioned to explore decision-making in a bottom-up way. Past comparative studies, for example, often assumed an inherent uniqueness to human behaviour: viewing differences between the predictions of economic models and animal behaviour as reliable differences between the biology of humans and other animals; not as failings of the model to account for other, external factors. Similarly, while economists have long asserted that utility maximisation models serve as 'as-if' models of decision-making (M. Friedman, 1953; Samuelson, 1937), neuroeconomics readily adopted utility's framework as a tool to locate decision-making signals in the brain (Fox & Poldrack, 2009; Wu & Glimcher, 2018). Where neuroeconomists could have asked *how* decisions were being made, they focused on *where*. This has made the current interpretations of neural activity particularly problematic since, other than providing a good fit to existing data, the is little evidence that 'as-if' utility models accurately predict out-of-sample data (D. Friedman et al., 2014; Volz & Gigerenzer, 2014) – let alone neural activity (O'Doherty, 2014). Expected Utility Theory is not and should not be seen as a theory about the brain – we should instead seek to explain EUT using what we know about the brain.

Likely a symptom of science's emphasis on 'significant' differences, the multidisciplinary-yet-top-down approach to studying decision-making has thus focused mainly on correlating differences between animals and brains with the differences in behaviour recognised by economics. The work in this thesis, in contrasts, looked for shared, bottom-up features in behaviour. In doing so, my work positioned itself closer to neuroethological thought than that of neuroeconomics: using biologically recognized and plausible mechanisms to explain divergent patterns of behaviour (rather than the other way around).

Through three separate experiments, that all shared a setup and stimuli, I showed that different 'irrational' preferences could result from single, stable mechanisms of adaptative preferences. That preferences adapt to the range of rewards that an animal could expect, for example, led to the same patterns of choice described by Prospect Theory's value function: animals were risk-seeking for rewards that fell below 'mean' expectations and risk-averse for those that fell above (Kahneman & Tversky, 1979). Similarly, the shift from concave (almost linear) utilities to s-shaped ones – when moving from riskless to risky choices – fell well in line with the idea that decision-makers perceive and indeed adapt their preferences to fit their expectations. Here the comparison between gamble and safe rewards might have anchored expectations, or indeed painted a gamble's lower outcome as a loss (Blanchard, Wilke, et al., 2014; Koszegi & Rabin, 2007; Wenner, 2015). Chapter 4, though, provided the most compelling example of how bottom-up processes could influence top-down metrics: the same win-stay/lose-shift algorithm, when applied to different reward schedules, led to vastly different and incompatible probability distortions.

Crucially, the reinterpretation of stable preferences into adaptive ones is supported by a growing body of work that suggests that what is consistent between animals is not, in fact, risk attitudes, but rather the way that risk attitudes adapt to contextual modulators (Farashahi et al., 2018; Heilbronner, 2017). The idea that simple, biologically-valid algorithms can support top-down observations is also one that decision theory has recently embraced (Mobbs et al., 2018; Volz & Gigerenzer, 2014; Wu & Glimcher, 2018). Ecologists, for example, have looked at several alternative theories to explain the ‘irrational’ behaviour that Prospect Theory predicts (Houston et al., 2014; McDermott et al., 2008; McNamara et al., 2014) – and a growing number of psychologists and neuroscientists go down the path of heuristics to interpret both brain signals and behavioural measurements (Khader et al., 2011; Mallpress et al., 2015; Volz et al., 2010). Even in neuroeconomics, a field built on economics’ normative framework, there is renewed interest in re-focusing on the similarities between humans and other animals; examining these similarities through ethologically rather than economically-driven theories (Kalenscher & van Wingerden, 2011; Mobbs et al., 2018; Pearson et al., 2014). Where a reconciliation between bottom-up and top-down perspectives might have once seemed unattainable – the idea that adaptation represents the rational, ecological constant might finally bring every field of decision theory into sync.

5.3 The way forward…

It would be unfair to criticise decision theory’s focus on unrealistic models without also acknowledging the limitations of my own findings. In that regard, thought my work sets a precedent for interpreting decision-making using biologically-plausible algorithms, one must also recognize that it clearly lacks in terms of generalizability: my results rest on the behaviour of but 2 or 3 monkeys, monkeys who have spent their whole lives in captive, laboratory housing. Without also studying larger or wilder groups of rhesus macaques, there is no way to know if Tigger, Trident, and Ugo’s behaviour was truly representative of their species as a whole (but a promising line of research aims to resolve this, see De Petrillo & Rosati, 2019). Likewise, the artificial and solitary situation in which these monkeys made choices is nothing like the natural and social environment in which they evolved (Barrett & Henzi, 2005; Hammerstein & Stevens, 2012; Seed & Tomasello, 2010). It would therefore be interesting to see how more naturalistic conditions impact behaviour – particularly given the recent finding that monkeys become risk-averse when allowed to freely move an forage (Eisenreich et al., 2019).

Because of the above, instead of providing conclusive answers the findings from this thesis open a Pandora's box of new questions (particularly when it comes to the interpretation of behaviours once recognized as static). For example, a set of questions emerges with regards to the divergent risk attitudes that captive primates have been shown to exhibit (see **Chapter 1**). How could adaptation reconcile macaque monkeys’ shift from risk-seeking behaviour in one set of studies (McCoy & Platt, 2005;

O'Neill & Schultz, 2010; Smith et al., 2017; Stauffer et al., 2015; E. R. Xu & Kralik, 2014) to risk-averse behaviour in another (Genest et al., 2016; Yamada et al., 2013). Likewise, rather than attributing the risk-averse behaviour of Bonobos and risk-seeking behaviour of Chimpanzees to their evolutionary history (Heilbronner, Rosati, Stevens, Hare, & Hauser, 2008; Krupenye et al., 2015), the difference might simply reflect preferences that have adapted to different tasks or environments contingencies: nature versus nurture. Adaptation to different ecological niches might then explain why bonobos exhibit risk-seeking behaviour when presented with choices in a different context (Haun et al., 2011). That primate preferences are flexible is by no means surprising. In fact, from the perspective of evolution, what has led primates to have such tremendous success is their inherent ability to adapt to different environments (Strier, 2011). What is surprising rather is that we biologists have been so quick to classify differences as static and stable.

Another important question that also remains to be studied is whether failing to win has the same effect on preferences as losing (or in the case of our monkeys, losing juice; Paglieri et al., 2014). In other words, do monkeys perceive the lower magnitude of a gamble the same way that an actual loss would be perceived? This is particularly relevant but difficult to study since primates have not been known to store rewards in the way humans 'store' money (Santos & Rosati, 2015). A promising and rather novel way around this would be to use digital tokens that accumulate (or vanish) based on a monkeys' decisions (a design that is slowly making its way into the literature; see Heilbronner, 2017; Paglieri et al., 2014). Under this design, macaques make choices that can add tokens to a 'token bank', or that remove tokens directly from it – the monkeys receiving a juice payoff only when the digital tokens in the bank reach a set number (Seo et al., 2014; Seo & Lee, 2009; Taswell et al., 2018). While this new design still plays on a monkey's later expectations, i.e., monkeys never lose more than they had when they first started the experiment with, it comes much closer to true loss than failing to win (Zakrzewski et al., 2014). It would be interesting to explore how preferences adapt to true, certain losses versus risky losses.

Finally, with behavioural flexibility being such a defining feature of primates (particularly humans and macaques), one might wonder if what differentiates us from other species is the way in which our brain and its neurons adapt their computations to different contexts. Recent evidence points to exactly that: full and partial range adaptation (i.e., refining one's expectations within a broader range of possibilities) of prefrontal neurons' firing rates is likely to underlie our ability to differentiate between contexts, to optimize our decisions in each of them, and to shape our understanding of how different situations fit together (Bavard et al., 2018; Conen & Padoa-Schioppa, 2019; Rustichini et al., 2017; Schuck et al., 2016; Wilson et al., 2014). If we are to understand the exact ways that adaptive brain signals contribute to behaviour, however, we need models that can capture and explain these processes. In short, it is not enough to study adaptive processes using static models. To understand dynamic processes, we need dynamic variables. To understand adaptive choices, we need adaptive economics.

5.4 Conclusion

The research presented in this thesis forms the basis for a reinterpretation of irrational choices a rational optimization of our limited brain - one that recognizes both the neuroethology of decision-making and the economic patterns of choice that have now been recognized for decades. My experimental results established that macaques' preferences are flexible, dynamic, and shaped by the reward statistics of the task at hand; it is likely that this ecologically rational mechanism is what leads to some of the more economically 'irrational' choices that we make.

Postface

~ All models are wrong, but adaptive models are better ~

A new generation of economic choice models: predictions and hypotheses

Stable preference metrics have been instrumental in neuroscientists' search for the brain areas that encode value and choice (Glimcher & Fehr, 2014). But where the idea of a 'utility' signal in the brain will eventually fail is in guiding our understanding of the computations that these brain areas perform. As advanced in this thesis, the fixed preference measurements from economic experiments might represent but a 'snapshot' or an average measure of the adaptive computations taking place throughout the brain (Balasubramani, Moreno-Bote, & Hayden, 2018; Ferrari-Toniolo, Bujold, & Schultz, 2019; Hunt & Hayden, 2017; Zimmermann, Glimcher, & Louie, 2018). These metrics, therefore, present an inherent limitation; they are indicative of what decisions are taking place where, but not how or why these decisions were made in the first place.

Though the work presented in this thesis focused on demonstrating that preference adaptation *did* occur, what follows serves mostly as an introduction to the advantages of switching from traditional economic models to a new generation of adaptive models. In particular, models that seek to explain dynamic preferences using neurons' fundamental properties (many exists, see e.g., Alos-Ferrer et al., 2018; Kontek & Lewandowski, 2018; Polanía, Woodford, & Ruff, 2019; Rustichini & Padoa-Schioppa, 2015; Soltani, de Martino, & Camerer, 2012). Of the many models that have cropped up in the last few years, Tymula & Glimcher's (2016) Expected Subjective Value Theory (ESVT) stands out as an elegantly simple but biologically-valid reformulation of expected utility theory – one that we can easily be fit to datasets that would otherwise be modelled using EUT or PT. ESVT remains a utility maximisation model, but it ascribes to its 'utility' function the same shape that governs neurons' firing rates (according to an already very successful model of brain activity; e.g., Louie, Grattan, & Glimcher, 2011; Hiroshi Yamada, Louie, Tymula, & Glimcher, 2018; Zimmermann et al., 2018). In essence, it keeps the variables traditionally used in economics (top-down perspective) but reinterprets the way decision-makers transform them in a way that makes biological sense (bottom-up). Just like visual neurons encode the brightness of an object relative to its background, ESVT assumes that decision-makers encode the value of different options relative to task-based expectations and context.

Just like expected utility theory, ESVT starts by defining the expected subjective value of a choice as:

$$ESV_t(x_{i,t}, p) = \sum_{i=1}^{n} p_i \times S(x_{i,t}) \qquad (6\text{-}1)$$

Where the expected subjective value of an option (at time *t*) is a function of a reward's outcome probabilities p_i and the subjective value of said outcomes $S(x_{i,t})$. The subjective value function, then, builds from a bottom-up interpretation of the way neurons encode information relative expectations:

$$S(x_{i,t}) = \frac{x_{i,t}^{\alpha}}{x_{i,t}^{\alpha}+Q_t^{\alpha}} \qquad (6\text{-}2)$$

In Tymula and Glimcher's words: "*The subjective value of [each]* $x_{i,t}$ *is divisively normalized by the reference point and itself"* (2016, p.5) - which leads to an s-shaped function that adapts to the decision-maker's expectations at time *t* (Q_t). The way ESVT is built differentiates it to traditional economic models with three explicit features: (i) preferences rely on inherently physical and comparable metric (firing rates), (ii) the function is bounded by the minimum and maximum firing rates that neurons exhibit, and (iii) just like neurons' firing rates, the value that ESVT assigns to rewards is normalized according to the distribution of possible outcomes (Louie et al., 2015; Wainwright et al., 2001; Webb et al., 2014). A fourth implicit feature is that, since the model's features rely on the properties of neurons, predictions from the model should apply equally to human and nonhuman animals. For anyone studying behaviour, ESVT offers a computationally simple yet powerful prediction of choice, risk-attitude, and – in its intended form – of neural firing rates.

Adaptive modelling in this thesis

To illustrate the uses and power that this new generation of models can have, let's revisit some of the results presented in this thesis. While my work looked at the repercussions of adaptation on static economic variables (i.e., utility and probability distortion), dynamic variables that adapt to the context of a task allow for a whole new set of falsifiable predictions.

I propose two additions to the ESVT model. The first is the inclusion of a random error term that allows us to bring ESVT under the random utility framework (see this thesis' **Methods interlogue**). This also brings the model more in line with the reality that the neural code is stochastic (Conen & Padoa-Schioppa, 2015; Kable & Glimcher, 2009; Kurtz-David et al., 2019; Shadlen & Newsome, 2001; Webb, 2018; S. D. Wilke & Eurich, 2002):

$$V(x_{i,t}, p_{i,t}) = p_i U(x_{i,t}) + \epsilon \qquad (6\text{-}3)$$

The second addition is the incorporation of reinforcement learning to account for dynamic adaptation of the decision-maker's reference-point (one that adapts to both the observed and obtained rewards; see Burke, Baddeley, Tobler, & Schultz, 2016):

$$Q_{t+1} = Q_t + \sum \delta_{EV}(x_t - Q_t) + \delta_{outcome}(R_t - Q_t) \qquad (6\text{-}4)$$

In this form, the expectations for future trials adapt (with strength δ) to the differences between past expectations and observed/obtained rewards. We define the starting expectation starts as the mean of the first observed variables - the expectation then adapts based on the predictions errors on all observed ($\sum \delta_{EV}(x_t - Q_t)$) and obtained rewards ($\delta_{outcome}(R_t - Q_t)$).

Applying this model to the case of probability distortion reversal offers the most clear-cut example of where and how adaptive choice models offer the most significant insight. In chapter 4, I suggested that preferences adapted to the wins and losses of past gambles. I then validated this by presenting the monkeys with MIXED or REPEATED sequences of gambles from which I retrieved two distinct patterns (or models) of ‘average’ risk attitudes. We can make behavioural predictions that are much more intricate using a model of adaptive economic behaviour – one that actually predicts moment-to-moment behaviour. If we think that the value of gambles changes relative the value of safe magnitudes, all we need to do is look at how a single instance of the ESVT’s model changes within and between the different sequences (Fig. 6-1). If we want to validate these predictions, we can look to find neurons that emulate this adaptive activity, or behaviourally, compare the model’s predictions of choice accuracy and risk attitude with the monkey’s real behaviour (Fig. 6-2).

In line with ecological rationality’s idea of the adaptive toolbox, this set algorithm can also be used hierarchically (where one normalization functions feeds into another), or to compute the relative value of different decision variables (think time discounting, uncertainty, or bundle-type rewards). There is evidence that the brain normalizes different signals in different regions and on different timescales (Hunt & Hayden, 2017; Murray et al., 2014; Zimmermann et al., 2018), ESVT-type models are simple models that can account for this.

Similarly, in a less radical/more incremental way, we can easily add variables to the model that allow us to investigate how monkey’s ‘starting’ expectations change as a result of context-specific adaptation. Looking at the results of Chapter 2, for example, an ESVT model predicts more than just ‘different’ patterns of utility for different ranges: it suggests that monkeys start with broad expectations that match past experiences, but that they, over the course of a task, refine their expectations to match the present instalment of the task (Fig. 6-3). Future work will undoubtedly explore the predictions that these biologically-valid models make – both at the behavioural and neural levels.

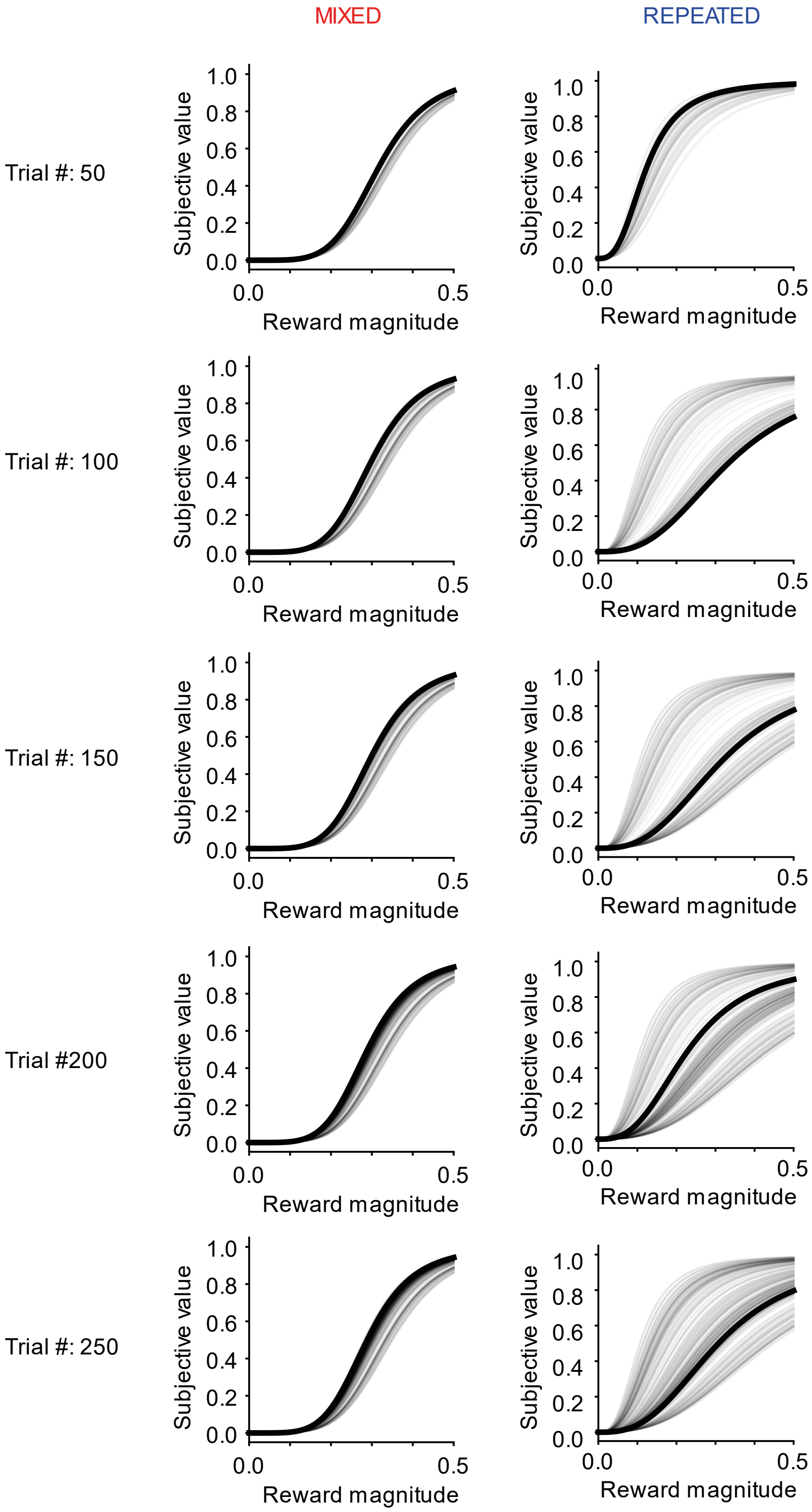


Figure 6-1 | Adaptive expected subjective values in MIXED or REPEATED sequences. The ‘utility’ functions (subjective value) adapts to past experienced and observed outcomes. In MIXED sequences, the utility fluctuates only a little since any reward can occur at any time (no skew in the options available). In the REPEATED sequences, however, prolonged exposure to a single gamble skews the utility function in the direction of that gamble’s value. The utility function adapts to each specific gamble. This happens because monkeys’ observing and receiving the rewards attributed to unique gambles drive their expectations towards that gamble’s expected value (through reinforcement learning). Crucially, the model predicts a reversal of risk preferences without the need for any probability distortion.

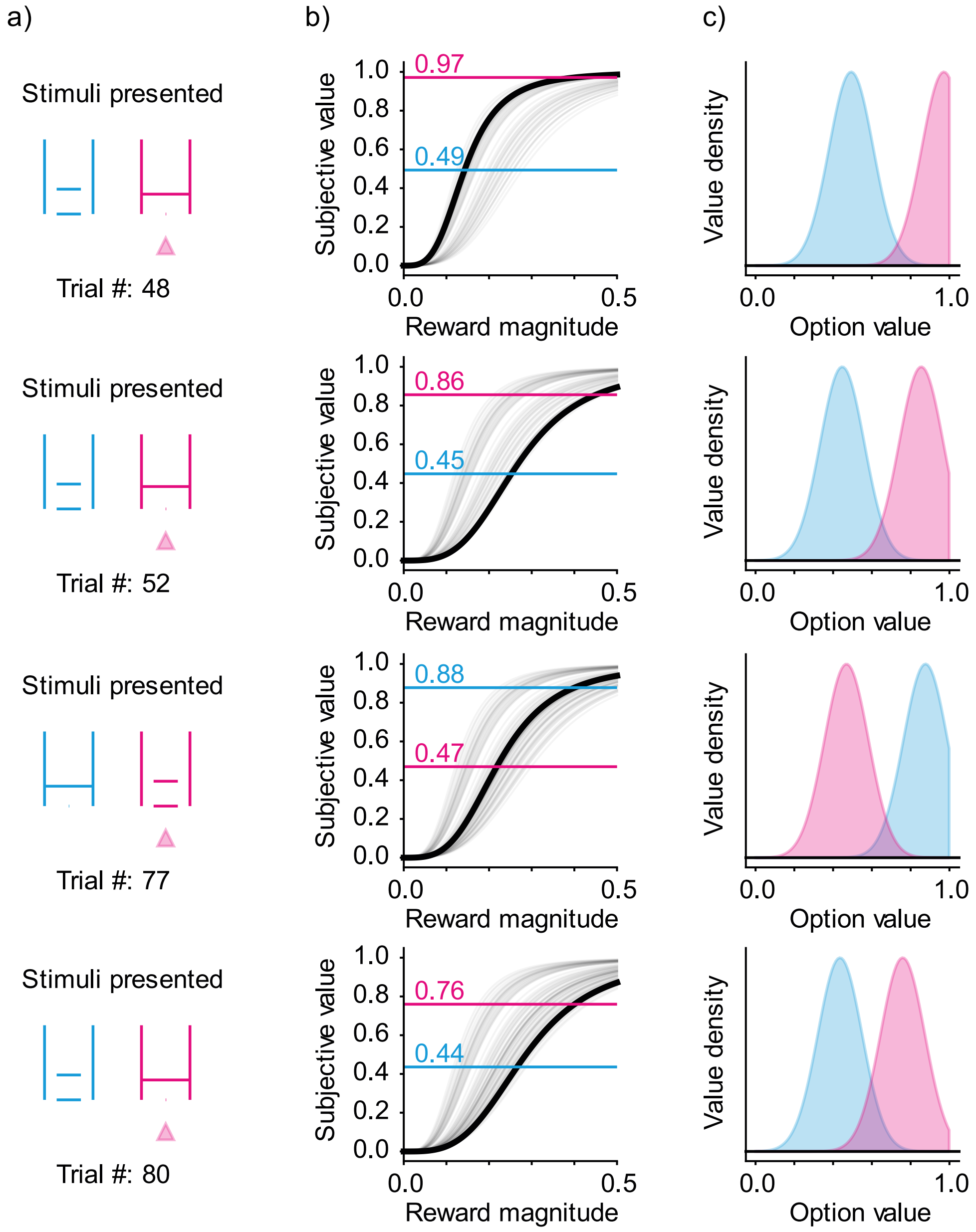


Figure 6-2 | Predicting choice stochasticity through random utility modelling (REPEATED sequence). The value function in the ESVT model, when combined with the random utility framework, allows for predictions on the separability of different rewards' utilities given a monkeys' past experiences in the task. a) choices between equiprobable gambles and safe reward. The ESVT model treats the magnitudes as subjective, and the probabilities as objective (for simplicity). The arrow under the options points to the option that was picked by the animal. b) ESVT's value function adapts throughout the task. The value function shifts to the right as the blocks of repeated gambles progressively shift expectations. This changes the value of the two outcomes relative one another and leads to different predictions on the reliability with which monkeys will pick the higher-value reward, and vis-à-vis the monkey's risk attitude. c) The overlap in random utilities distributions changes with adaptation. The more two distributions overlap, the more stochastic the monkeys' decisions will be. The left option is represented at every level of the model (panels a, b, c) in blue, the right option in pink (magenta).

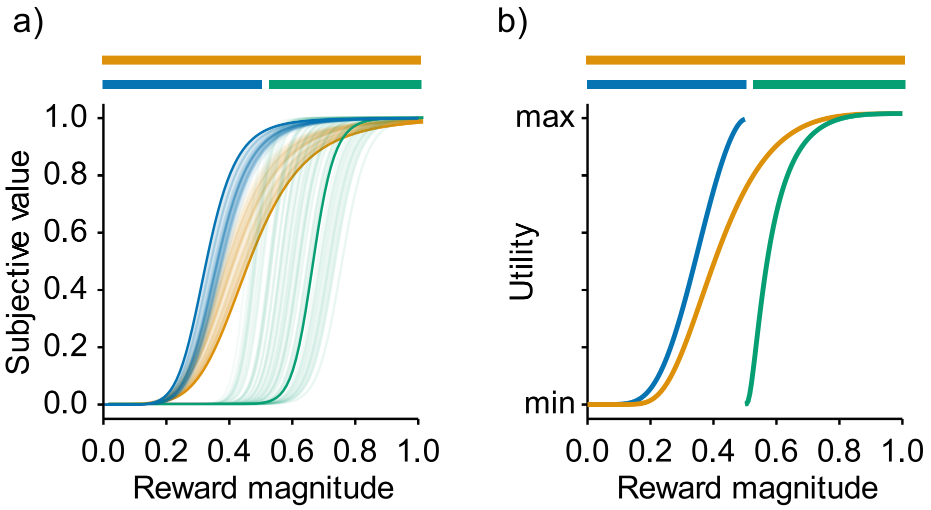


Figure 6-3 | Range-adaptation refines preferences within a single range. a) ESVT value functions suggest monkeys go into high-range sessions with low expectations. While the value function in both the low-range (blue) and the full-range (yellow) predict that the monkey's expectations stay relatively constant, the predictions from the high-range value function (green) suggests that monkeys' broader expectations get refined as the session progresses. Crucially, these predictions can be tested since risk-preferences still depend on the value function's curvature at different magnitudes. b) EUT and PT values only capture 'average' behaviour. Where ESVT predicts that monkey's preferences adapt progressively throughout the task, traditional economic models cannot make a distinction between where expectations at the beginning of an experimental sessions and expectations at the end.

Bibliography

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science*, *46*(11), 1497–1512.

Abdellaoui, M., Barrios, C., & Wakker, P. P. (2007). Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory. *Journal of Econometrics*, *138*(1), 356–378.

Abdellaoui, M., Bleichrodt, H., L'Haridon, O., & Paraschiv, C. (2013). Is There One Unifying Concept of Utility?An Experimental Comparison of Utility Under Risk and Utility Over Time. *Management Science*, *59*(9), 2153–2169.

Abdellaoui, M., L'Haridon, O., & Paraschiv, C. (2011). Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications? *Management Science*, *57*(10), 1879–1895.

Abrahamyan, A., Silva, L. L., Dakin, S. C., Carandini, M., & Gardner, J. L. (2016). Adaptable history biases in human perceptual decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(25), E3548–E3557.

Adams, G. K., Watson, K. K., Pearson, J., & Platt, M. L. (2012). Neuroethology of decision-making. *Current Opinion in Neurobiology*, *22*(6), 982–989.

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica*, *21*(4), 503.

Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2016). Reference-Dependent Preferences: Evidence from Marathon Runners. *Management Science*, *63*(6), 1657–1672.

Alos-Ferrer, C., Fehr, E., Netzer, N., Alós-ferrer, C., Fehr, E., & Netzer, N. (2018). *Time will tell - Recovering Preferences when Choices are Noisy*. *July*.

Andersen, S., Harrison, G. W., Lau, M. I., Rutström, E. E., Rutstrom, E. E., & Rutstrom, Elisabet, E. (2011). Eliciting Risk and Time Preferences. *Econometrica*, *76*(3), 583–618.

Andreoni, B. J., & Sprenger, C. (2012). *Estimating Time Preferences from Convex Budgets*. *102*(7), 3333–3356.

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. Harper Collins.

Arkes, H. R., Hirshleifer, D., Jiang, D., & Lim, S. (2008). Reference point adaptation: Tests in the domain of security trading. *Organizational Behavior and Human Decision Processes*, *105*(1), 67–81.

Arkes, H. R., Hirshleifer, D., Jiang, D., & Lim, S. S. (2010). A cross-cultural study of reference point adaptation: Evidence from China, Korea, and the US. *Organizational Behavior and Human Decision Processes*, *112*(2), 99–111.

Avugos, S., Köppen, J., Czienskowski, U., Raab, M., & Bar-Eli, M. (2013). The " hot hand" reconsidered: A meta-analytic approach. In *Psychology of Sport and Exercise* (Vol. 14, Issue 1, pp. 21–27).

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory and Cognition*, *32*(8), 1369–1378.

Baillon, A., Bleichrodt, H., & Spinu, V. (2015). *Searching for the Reference Point* (Issue May).

Balasubramani, P. P., Moreno-Bote, R., & Hayden, B. Y. (2018). Using a Simple Neural Network to Delineate Some Principles of Distributed Economic Choice. *Frontiers in Computational Neuroscience*, *12*(March), 1–13.

Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. (1992). Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, *25*(2), 390–397.

Bar-Eli, M., Avugos, S., & Raab, M. (2006). Twenty years of "hot hand" research: Review and critique. In *Psychology of Sport and Exercise* (Vol. 7, Issue 6, pp. 525–553).

Barberis, N. (2012). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Ssrn*, *27*(1), 173–196.

Barrett, L., & Henzi, P. (2005). The social nature of primate cognition. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1575), 1865–1875.

Barron, G., & Erev, I. (2003). Small Feedback-based Decisions and Their Limited Correspondence to Description-based Decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233.

Barron, H. F., Winterfeldt, D. von, & Fischer, G. W. (1984). Empirical and theoretical relationships between value and utility functions. *Acta Psychologica*, *56*(1–3), 233–244.

Barss, P. (1984). Injuries due to falling coconuts. *Journal of Trauma - Injury, Infection and Critical Care*.

Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature Communications*, *9*(1).

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141.

Becker, G. M., Degroot, M. H., & Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Science*.

Bernoulli, D. D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23.

Blanchard, T. C., & Hayden, B. Y. (2014). Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *Journal of Neuroscience*.

Blanchard, T. C., Pearson, J. M., & Hayden, B. Y. (2013). Postreward delays and systematic biases in measures of animal temporal discounting. *Proceedings of the National Academy of Sciences*, *110*(38), 15491–15496.

Blanchard, T. C., Wilke, A., & Hayden, B. Y. (2014). Hot-hand bias in rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, *40*(3), 280–286.

Blanchard, T. C., Wolfe, L. S., Vlaev, I., Winston, J. S., & Hayden, B. Y. (2014). Biases in preferences for sequences of outcomes in monkeys. *Cognition*, *130*(3), 289–299.

Block, H. D., & Marschak, J. (1959). Random Orderings and Stochastic Theories of Responses. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*.

Boggs, P. T., & Rogers, J. E. (2012). *Orthogonal distance regression*.

Bossaerts, P., & Murawski, C. (2015). From behavioural economics to neuroeconomics to decision neuroscience: The ascent of biology in research on human decision making. *Current Opinion in Behavioral Sciences*, *5*, 37–42.

Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. In *Robustness in Statistics* (pp. 201–236). Elsevier.

Boyden, E. S., Borel, M., Schultz, W., Yang, A., Stauffer, W. R., Paulsen, O., & Lak, A. (2016). Dopamine Neuron-Specific Optogenetic Stimulation in Rhesus Macaques. *Cell*, *166*(6), 1564-1571.e6.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky Choice With Heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, and Willemsen (2008), and Rieger and Wang (2008). *Psychological Review*, *115*(1), 281–289.

Brandstätter, E., & Gussmack, M. (2013). The Cognitive Processes Underlying Risky Choice. *Journal of Behavioral Decision Making*, *26*(2), 185–197.

Britten, K. H., & Heuer, H. W. (1999). Spatial summation in the receptive fields of MT neurons. *Journal of Neuroscience*.

Brosnan, S. F., Wilson, B. J., & Beran, M. J. (2012). Old world monkeys are more similar to humans than new world monkeys when playing a coordination game. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1733), 1522–1530.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica*, *78*(4), 1375–1412.

Brunswik, E. (1956). Perception and the representative design of psychological experiments, 2nd ed. In *Perception and the representative design of psychological experiments, 2nd ed.*

Burke, C. J., Baddeley, M., Tobler, P. N., & Schultz, W. (2016). Partial Adaptation of Obtained and Observed Value Signals Preserves Information about Gains and Losses. *Journal of Neuroscience*, *36*(39), 10016–10025.

Burke, C. J., Soutschek, A., Weber, S., Raja Beharelle, A., Fehr, E., Haker, H., & Tobler, P. N. (2018). Dopamine Receptor-Specific Contributions to the Computation of Value. *Neuropsychopharmacology*, *43*(6), 1415–1424.

Burns, B. D. (2004). Heuristics as beliefs and as behaviors: The adaptiveness of the “hot hand.” *Cognitive Psychology*, *48*(3), 295–331.

Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: “Gambler’s fallacy” versus “hot hand.” *Psychonomic Bulletin and Review*, *11*(1), 179–184.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*.

Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, *2*(1), 61–104.

Camerer, C. F. (1998). Bounded rationality in individual decision making. *Experimental Economics*, *1*(2), 163–183.

Camerer, C. F., Babcock, L., Loewenstein, G., & Thaler, R. (2002). Labor Supply of New York City Cabdrivers: One Day at a Time. *The Quarterly Journal of Economics*, *112*(2), 407–441.

Camerer, C. F., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How Neuroscience Can Inform Economics. *Journal of Economic Literature*, *43*(1), 9–64.

Camerer, C. F., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*(4), 325–370.

Camilleri, A. R., & Newell, B. R. (2013). The long and short of it: Closing the description-experience "gap" by taking the long-run view. *Cognition*, *126*(1), 54–71.

Capitanio, J. P., & Emborg, M. E. (2008). Contributions of non-human primates to neuroscience research. *The Lancet*, *371*(9618), 1126–1135.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience*.

Cartwright, D., & Festinger, L. (1943). A quantitative theory of decision. *Psychological Review*, *50*(6), 595–621.

Castel, A. D., Rossi, A. D., & McGillivray, S. (2012). Beliefs About the "Hot Hand" in Basketball Across the Adult Life Span. *Psychology and Aging*, *27*(3), 601–605.

Chase, V. M., Hertwig, R., & Gigerenzer, G. (1998). Visions of rationality. In *Trends in Cognitive Sciences* (Vol. 2, Issue 6, pp. 206–214).

Chen, K. M., Lakshminarayanan, V. R., & Santos, L. (2006). How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior. *Journal of Political Economy*, *114*(3), 517–537.

Chen, X., & Stuphorn, V. (2018). Inactivation of Medial Frontal Cortex Changes Risk Preference. *Current Biology*, *28*(19), 3114-3122.e4.

Cheung, S. L. (2015). *Eliciting utility curvature and time preference*. *January*.

Cheung, S. L. (2016). Recent developments in the experimental elicitation of time preference. *Journal of Behavioral and Experimental Finance*, *11*, 1–8.

Choi, S. J., & Pritchard, A. C. (2003). Behavioral Economics and the SEC. *Stanford Law Review*, *56*(1), 1–73.

Chung, H.-K., Glimcher, P., & Tymula, A. (2019). An Experimental Comparison of Risky and Riskless Choice—Limitations of Prospect Theory and Expected Utility Theory. *American Economic Journal: Microeconomics*, *11*(3), 34–67.

Conen, K. E., & Padoa-Schioppa, C. (2015). Neuronal variability in orbitofrontal cortex during economic decisions. *Journal of Neurophysiology*, *114*(3), 1367–1381.

Conen, K. E., & Padoa-Schioppa, C. (2019). Partial Adaptation to the Value Range in the Macaque Orbitofrontal Cortex. *The Journal of Neuroscience*, 2279–18.

Cowling, T. G. (1955). The Letters of Johann Bernoulli. *Nature*, *176*(1187), 1187–1187.

Cox, K. M., & Kable, J. W. (2014). BOLD Subjective Value Signals Exhibit Robust Range Adaptation. *The Journal of Neuroscience*, *34*(49), 16533–16543.

Crawford, V. P., & Meng, J. (2011). New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences. *American Economic Review*, *101*(5), 1912–1932.

Croson, R., & Sundali, J. (2005). The Gambler's Fallacy and the Hot Hand: Empirical Data from Casinos. In *The Journal of Risk and Uncertainty* (Vol. 30, Issue 3).

Csapo, P., Avugos, S., Raab, M., & Bar-Eli, M. (2015). How should "hot" players in basketball be defended? The use of fast-and-frugal heuristics by basketball coaches and players in response to streakiness. *Journal of Sports Sciences*, *33*(15), 1580–1588.

Daw, N. D., & Tobler, P. N. (2013). Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning. In *Neuroeconomics: Decision Making and the Brain: Second Edition*. Elsevier Inc.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases and rational decision-making in the human brain. *Science*, *313*(5787), 684–687.

De Petrillo, F., & Rosati, A. G. (2019). Rhesus macaques use probabilities to predict future events. *Evolution and Human Behavior*, *40*(5), 436–446.

Delquié, P., & Cillo, A. (2006). Disappointment without prior expectation: A unifying perspective on decision under risk. *Journal of Risk and Uncertainty*, *33*(3), 197–215.

Diederen, K. M. M. J., Spencer, T., Vestergaard, M. D. D., Fletcher, P. C. C., & Schultz, W. (2016). Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron*, *90*(5), 1127–1138.

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. In *Current Opinion in Neurobiology* (Vol. 22, Issue 6, pp. 1075–1081). Elsevier Ltd.

Eisenreich, B. R., Hayden, B. Y., & Zimmermann, J. (2019). Macaques are risk-averse in a freely moving foraging task. *Scientific Reports*, *9*(1), 1–12.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics*, *75*(4), 643–669.

Fairhall, A. L., Lewen, G. D., Bialek, W., & De Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, *412*(6849), 787–792.

Falk, R., & Konold, C. (1997). Making Sense of Randomness: Implicit Encoding as a Basis for Judgment. *Psychological Review*, *104*(2), 301–318.

Farashahi, S., Azab, H., Hayden, B., & Soltani, A. (2018). On the flexibility of basic risk attitudes in monkeys. *The Journal of Neuroscience*, *38*(18), 4383–4398.

Farquhar, P. H. (1984). State of the Art — Utility assessment methods. *Management Science*, *30*(11), 1283–1300.

Fennema, H., & Wakker, P. (1997). Original and cumulative prospect theory: a discussion of empirical differences. *Journal of Behavioral Decision Making*, *10*(1), 53–64.

Ferrari-Toniolo, S., Bujold, P. M., & Schultz, W. (2019). Probability distortion depends on choice sequence in rhesus monkeys. *The Journal of Neuroscience*, *39*(15), 1454–18.

Fiedler, K., & von Sydow, M. (2015). Heuristics and biases: Beyond Tversky and Kahneman (1974) judgment under uncertainty. *Cognitive Psychology: Revisting the Classical Study*, *January*, 146–161.

Fox, C. R., & Poldrack, R. A. (2009). Chapter 11 - Prospect theory and the brain. In *Neuroeconomics*.

Frey, R., Mata, R., & Hertwig, R. (2015). The role of cognitive abilities in decisions from experience: Age differences emerge as a function of choice set size. *Cognition*, *142*, 60–80.

Friedman, D., Isaac, R. M., James, D., & Sunder, S. (2014). *Risky curves: On the empirical failure of expected utility*. Routledge.

Friedman, M. (1953). Essays in Positive Economics. *University of Chicago Press*.

Genest, W., Stauffer, W. R., & Schultz, W. (2016). Utility functions predict variance and skewness risk preferences in monkeys. *Proceedings of the National Academy of Sciences*, *113*(30), 8402–8407.

Gigerenzer, G. (2010). Personal Reflections on Theory and Psychology. *Theory & Psychology*, *20*(6), 733–743.

Gigerenzer, G. (2016). Towards a Rational Theory of Heuristics. In *Minds, Models and Milieux*.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science*, *1*(1), 107–143.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, *62*(1), 451–482.

Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making*, *6*(1), 100–121.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A brunswikian theory of confidence. *Psychological Review*.

Gigerenzer, G., & Selten, R. (2002). *Bounded Rationality: The Adaptive Toolbox*. MIT Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*(3), 295–314.

Glimcher, P. W., Dorris, M. C., & Bayer, H. M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games and Economic Behavior*, *52*(2), 213–256.

Glimcher, P. W., & Fehr, E. (2014). Introduction: A Brief History of Neuroeconomics. *Neuroeconomics (Second Edition)*, xvii–xxviii.

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition*, *133*(3), 641–666.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*(1), 75–90.

Gonzalez, R., & Wu, G. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology*, *166*(1), 129–166.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112.

Hammack, T., Cooper, J., Flach, J. M., & Houpt, J. (2017). Toward an Ecological Theory of Rationality: Debunking the Hot Hand “Illusion.” *Ecological Psychology*, *29*(1), 35–53.

Hammerstein, P., & Hagen, E. H. (2005). The second wave of evolutionary economics in biology. *Trends in Ecology and Evolution*, *20*(11), 604–609.

Hammerstein, P., & Stevens, J. R. (2012). Six Reasons for Invoking Evolution in Decision Theory. In *Evolution and the Mechanisms of Decision Making* (pp. 1–18).

Harless, D. W., & Camerer, C. F. (1994). The Predictive Utility of Generalized Expected Utility Theories. *Econometrica*, *62*(6), 1251.

Haun, D. B. M., Nawroth, C., & Call, J. (2011). Great Apes’ risk-taking strategies in a decision making task. *PLoS ONE*, *6*(12), 1–6.

Hayden, B. Y. (2018). Economic choice: the foraging perspective. *Current Opinion in Behavioral Sciences*, *24*, 1–6.

Hayden, B. Y., Heilbronner, S. R., Nair, A. C., & Platt, M. L. (2008). Cognitive influences on risk-seeking by rhesus macaques. *Judgment and Decision Making*, *3*(5), 389–395.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–939.

Hayden, B. Y., & Platt, M. L. (2007). Temporal Discounting Predicts Risk Sensitivity in Rhesus Macaques. *Current Biology*, *17*(1), 49–53.

Hayden, B. Y., & Walton, M. E. (2014). Neuroscience of foraging. *Frontiers in Neuroscience*, *8*, 81.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*(2), 181–197.

Heilbronner, S. R. (2017). Modeling risky decision-making in nonhuman animals: shared core features. *Current Opinion in Behavioral Sciences*, *16*(3), 23–29.

Heilbronner, S. R., & Hayden, B. Y. (2013). Contextual factors explain risk-seeking preferences in rhesus monkeys. *Frontiers in Neuroscience*, *7*(7 FEB), 1–7.

Heilbronner, S. R., & Hayden, B. Y. (2016). The description-experience gap in risky choice in nonhuman primates. *Psychonomic Bulletin and Review*, *23*(2), 593–600.

Heilbronner, S. R., Rosati, A. G., Stevens, J. R., Hare, B., & Hauser, M. D. (2008). A fruit in the hand or two in the bush? Divergent risk preferences in chimpanzees and bonobos. *Biology Letters*, *4*(3), 246–249.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization*.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, *15*(8), 534–539.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*(2), 225–237.

Hertwig, R., Wulff, D. U., & Mata, R. (2018). Three gaps and what they may mean for risk preference. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1766), 20180140.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental Economics*, *8*(4 SPEC. ISS.), 325–345.

Hey, J. D., & Strazzera, E. (1989). Estimation of indifference curves in the Marschak-Machina triangle a direct test of the “fanning out” hypothesis. *Journal of Behavioral Decision Making*, *2*(4), 239–260.

Hoffman, E., McCabe, K. A., & Smith, V. L. (1998). Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology. *Economic Inquiry*, *36*(3), 335–352.

Houston, A. I., Fawcett, T. W., Mallpress, D. E. W., & McNamara, J. M. (2014). Clarifying the relationship between prospect theory and risk-sensitive foraging theory. *Evolution and Human Behavior*, *35*(6), 502–507.

Howard, J. D., & Kahnt, T. (2017). Identity-Specific Reward Representations in Orbitofrontal Cortex Are Modulated by Selective Devaluation. *The Journal of Neuroscience*, *37*(10), 2627–2638.

Hsu, M., Krajbich, I., Zhao, C., & Camerer, C. F. (2009). Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities. *Journal of Neuroscience*, *29*(7), 2231–2237.

Huk, A. C. (2005). Neural Activity in Macaque Parietal Cortex Reflects Temporal Integration of Visual Motion Signals during Perceptual Decision Making. *Journal of Neuroscience*, *25*(45), 10420–10436.

Hunt, L. T., & Hayden, B. Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews Neuroscience*, *18*(3), 172–182.

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*.

Hunter, L. E., & Gershman, S. J. (2018). Reference-dependent preferences arise from structure learning. *BioRxiv*, 252692.

Hwang, J., Kim, S., & Lee, D. (2009). Temporal discounting and inter-temporal choice in rhesus monkeys. *Frontiers in Behavioral Neuroscience*, *3*(June), 1–13.

Jones, C. B. (2005). Behavioral Flexibility: Interpretations and Prospects. In *Behavioral Flexibility in Primates: Causes and Consequences* (pp. 123–138). Springer US.

Kable, J. W., & Glimcher, P. W. (2009). The Neurobiology of Decision: Consensus and Controversy. *Neuron*, *63*(6), 733–745.

Kacelnik, A. (2012). Meanings of rationality. In *Rational Animals?*

Kahneman, D. (2011). *Thinking, fast and slow*. NY: Farrar, Straus and Giroux.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, *98*(6), 1325–1348.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291.

Kalenscher, T., & van Wingerden, M. (2011). Why we should use animals to study economic decision making - A perspective. *Frontiers in Neuroscience*, *5*(JUN), 1–11.

Kalenscher, T., & Wingerden, M. Van. (2013). *Behavioral and neuroscientific analysis of economic decision making in animals* (Frontiers). Frontiers.

Keeney, R. L., & Raiffa, H. (1993). Decisions with Multiple Objectives: Preferences and Value Trade-Offs. In *Cambridge: Cambridge University Press* (Vol. 9, Issue 7). Cambridge University Press.

Kellen, D., Pachur, T., & Hertwig, R. (2016). How (in)variant are subjective representations of described and experienced risk and rewards? *Cognition*, *157*, 126–138.

Khader, P. H., Pachur, T., Meier, S., Bien, S., Jost, K., & Rösler, F. (2011). *Memory-based Decision-making with Heuristics: Evidence for a Controlled Activation of Memory Representations*. www.psy.uwa.edu.au/MRCDataBase/

Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., & Loewenstein, G. (2008). Neural Antecedents of the Endowment Effect. *Neuron*, *58*(5), 814–822.

Kobayashi, S., & Schultz, W. (2008). Influence of Reward Delays on Responses of Dopamine Neurons. *Journal of Neuroscience*, *28*(31), 7837–7846.

Kobayashi, Shunsuke, Pinto de Carvalho, O., & Schultz, W. (2010). Adaptation of Reward Sensitivity in Orbitofrontal Neurons. *Journal of Neuroscience*, *30*(2), 534–544.

Kolling, N., Wittmann, M., & Rushworth, M. F. S. (2014). Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron*, *81*(5), 1190–1202.

Kong, Q., Granic, G. D., Lambert, N. S., & Teo, C. P. (2019). Judgment Error in Lottery Play: When the Hot Hand Meets the Gambler's Fallacy. *Management Science*.

Konishi, M. (2010). Neuroethology: What is it? In *Encyclopedia of Animal Behavior* (pp. 562–565). Elsevier.

Kontek, K., & Lewandowski, M. (2018). Range-Dependent Utility. *Management Science*, *64*(6), 2812–2832.

Koszegi, B., & Rabin, M. (2007). Reference-Dependent Risk Attitudes. *American Economic Review*, *97*(4), 1047–1073.

Köszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165.

Kotz, S., & Dorp, J. R. van. (2010). Beyond Beta - Other Continuous Families of Distributions with Bounded Support and Applications. In *Beyond Beta - Other Continuous Families of Distributions with Bounded Support and Applications*.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, *93*(3), 480–490.

Krupenye, C., Rosati, A. G., & Hare, B. (2015). Bonobos and chimpanzees exhibit human-like framing effects. *Biology Letters*, *11*(2), 10–13.

Kuchibhotla, K. V., Gill, J. V., Lindsay, G. W., Papadoyannis, E. S., Field, R. E., Sten, T. A. H., Miller, K. D., & Froemke, R. C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. *Nature Neuroscience*, *20*(1), 62–71.

Kumar, S., & Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature*, *392*(6679), 917–920.

Kurtz-David, V., Persitz, D., Webb, R., & Levy, D. J. (2019). The neural computation of inconsistent choice behavior. *Nature Communications*, *10*(1), 1583.

Kvitsiani, D., Ranade, S., Hangya, B., Taniguchi, H., Huang, J. Z., & Kepecs, A. (2013). Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature*.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., Cheng -, H. M., Brady, T. J., & Rosen, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(12), 5675–5679.

Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences*, *111*(6), 2343–2348.

Lakshminarayanan, V. R., Chen, K. M., & Santos, L. R. (2011). The evolution of decision-making under risk: Framing effects in monkey risk preferences. *Journal of Experimental Social Psychology*, *47*(3), 689–693.

Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice. A parametric examination. *Journal of Economic Behavior and Organization*, *17*(3), 377–400.

Lau, B., & Glimcher, P. W. (2005). Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys. *Journal of the Experimental Analysis of Behavior*, *84*(3), 555–579.

Laughlin, S. (1981). A Simple Coding Procedure Enhances a Neuron's Information Capacity. *Zeitschrift Für Naturforschung*, *36*(9–10), 910–912.

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, *116*(2), 286–295.

Levitt, S. D., & Lis, J. A. (2008). Economics: Homo economicus evolves. In *Science* (Vol. 319, Issue 5865, pp. 909–910).

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038.

LoFaro, T., Louie, K., Webb, R., & Glimcher, P. W. (2014). The Temporal Dynamics of Cortical Normalization Models of Decision-making. *Letters in Biomathematics*, *1*(2), 209–220.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, *39*(3–4), 641–648.

Loomes, G., & Sugden, R. (2006). Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *The Economic Journal*.

Lopez-Guzman, S., Konova, A. B., Louie, K., & Glimcher, P. W. (2018). Risk preferences impose a hidden distortion on measures of choice impulsivity. *PLOS ONE*, *13*(1), e0191357.

Louie, K., & De Martino, B. (2013). The Neurobiology of Context-Dependent Valuation and Choice. In *Neuroeconomics: Decision Making and the Brain: Second Edition*.

Louie, K., & Glimcher, P. W. (2012). Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, *1251*(1), 13–32.

Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: from biological to behavioral decision- making. *Current Opinion in Behavioral Sciences*, *5*(212), 91–99.

Louie, K., Grattan, L. E., & Glimcher, P. W. (2011). Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex. *Journal of Neuroscience*, *31*(29), 10627–10639.

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, *110*(15), 6139–6144.

Louie, K., LoFaro, T., Webb, R., & Glimcher, P. W. (2014). Dynamic Divisive Normalization Predicts Time-Varying Value Coding in Decision-Related Circuits. *Journal of Neuroscience*, *34*(48), 16046–16057.

Luce, R. D. (1958). A Probabilistic Theory of Utility. *Econometrica*.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*.

Machina, M. J. (1982). “Expected Utility” Analysis without the Independence Axiom. *Econometrica*, *50*(2), 277.

Machina, M. J. (1985). Stochastic Choice Functions Generated From Deterministic Preferences Over Lotteries. *The Economic Journal*, *95*(379), 575.

Machina, M. J. (1987). Choice Under Uncertainty: Problems Solved and Unsolved. *Journal of Economic Perspectives*, *1*(1), 121–154.

Mallpress, D. E. W., Fawcett, T. W., Houston, A. I., & McNamara, J. M. (2015). Risk attitudes in a changing environment: An evolutionary model of the fourfold pattern of risk preferences. *Psychological Review*, *122*(2), 364–375.

Marr, D., & Ullman, S. (1982). *VISION: A Computational Investigation into the Human Representation and Processing of Visual Information*.

Marschak, J. (1950). Rational Behavior, Uncertain Prospects, and Measurable Utility. *Econometrica*, *18*(3), 312.

Marschak, J. (1960). Binary Choice Constraints on Random Utility Indicators. In *Economic Information, Decision, and Prediction*.

Marshall, A. T., & Kirkpatrick, K. (2013). The effects of the previous outcome on probabilistic choice in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*(1), 24–38.

Marshall, A. T., & Kirkpatrick, K. (2015). Relative gains, losses, and reference points in probabilistic choice in rats. *PLoS ONE*, *10*(2), 1–33.

Masatlioglu, Y., & Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, *121*, 1–29.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience*, *8*(9), 1220–1227.

McDermott, R., Fowler, J. H., & Smirnov, O. (2008). On the evolutionary origin of prospect theory preferences. *Journal of Politics*, *70*(2), 335–350.

McFadden, D. (1974a). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*.

McFadden, D. (1974b). The measurement of urban travel demand. *Journal of Public Economics*.

McFadden, D. (2001). Economic Choices. *American Economic Review*, *91*(3), 351–378.

McFadden, D. (2005). Revealed stochastic preference: A synthesis. *Economic Theory*, *26*(2), 245–264.

McFadden, D. (2014). *Handbook of Choice Modelling - Google Books*. *1*, 1–2.

McFall, J. P. (2015). Rational, normative, descriptive, prescriptive, or choice behavior? The search for integrative metatheory of decision making. *Behavioral Development Bulletin*, *20*(1), 45–59.

McNamara, J. M., & Houston, A. I. (1986). The Common Currency for Behavioral Decisions. *The American Naturalist*, *127*(3), 358–378.

McNamara, J. M., & Houston, A. I. (2009). Integrating function and mechanism. In *Trends in Ecology and Evolution* (Vol. 24, Issue 12, pp. 670–675).

McNamara, J. M., Trimmer, P. C., & Houston, A. I. (2014). Natural selection can favour ‘irrational’ behaviour. *Biology Letters*, *10*(1), 20130935.

McNeil, B. J., Pauker, S. G., Sox, H. C., & Tversky, A. (1982). On the Elicitation of Preferences for Alternative Therapies. *New England Journal of Medicine*, *306*(21), 1259–1262.

Menard, S. (2011). Standards for Standardized Logistic Regression Coefficients. *Social Forces*, *89*(4), 1409–1428.

Miller, J. B., & Sanjurjo, A. (2018). Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers. *Econometrica*, *86*(6), 2019–2047.

Mishra, S. (2014). Decision-Making Under Risk: Integrating Perspectives From Biology, Economics, and Psychology. *Personality and Social Psychology Review*, *18*(3), 280–307.

Mitchell, A. S., Thiele, A., Petkov, C. I., Roberts, A., Robbins, T. W., Schultz, W., & Lemon, R. (2018). Continued need for non-human primate neuroscience research. *Current Biology*, *28*(20), R1186–R1187.

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: Insights from ethology. *Nature Reviews Neuroscience*, *19*(7), 419–427.

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. In *Neuron* (Vol. 36, Issue 2, pp. 265–284). Cell Press.

Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., & Wang, X. J. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience*, *17*(12), 1661–1663.

Neiman, T., & Loewenstein, Y. (2011). Reinforcement learning in professional basketball players. *Nature Communications*, *2*(1).

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, *341*(6237), 52–54.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review*, *109*(2), 330–357.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*(6432), 56–58.

O'Doherty, J. P. (2014). The problem with value. *Neuroscience and Biobehavioral Reviews*, *43*, 259–268.

O'Neill, M., & Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron*, *68*(4), 789–800.

O'Neill, M., & Schultz, W. (2018). Predictive coding of the statistical parameters of uncertain rewards by orbitofrontal neurons. *Behavioural Brain Research*, *355*(April), 90–94.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(13), 5951–5955.

Oliphant, T. E. (2007a). Python for Scientific Computing. *Computing in Science & Engineering*, *9*(3), 10–20.

Oliphant, T. E. (2007b). SciPy: Open source scientific tools for Python. *Computing in Science and Engineering*.

Ore, O. (1960). Pascal and the Invention of Probability Theory. *The American Mathematical Monthly*, *67*(5), 409.

Oskarsson, A. T., Van Boven, L., McClelland, G. H., & Hastie, R. (2009). What's Next? Judging Sequences of Binary Events. *Psychological Bulletin*, *135*(2), 262–285.

Pachur, T. (2011). The limited value of precise tests of the recognition heuristic. In *Judgment and Decision Making* (Vol. 6, Issue 5).

Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L. J., & Goldstein, D. G. (2011). The recognition heuristic: A review of theory and tests. *Frontiers in Psychology*, *2*(JUL), 1–14.

Padoa-Schioppa, C. (2009). Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex. *Journal of Neuroscience*, *29*(44), 14004–14014.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226.

Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*(1), 95–102.

Paglieri, F., Addessi, E., De Petrillo, F., Laviola, G., Mirolli, M., Parisi, D., Petrosino, G., Ventricelli, M., Zoratto, F., & Adriani, W. (2014). Nonhuman gamblers: lessons from rodents, primates, and robots. *Frontiers in Behavioral Neuroscience*, *8*(February), 1–18.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, *72*(6), 407–418.

Parducci, A. (2012). Contextual effects: a range-frequency analysis. In *Psychophysical Judgment and Measurement*.

Pastor-Bernier, A., Plott, C. R., & Schultz, W. (2017). Monkeys choose as if maximizing utility compatible with basic principles of revealed preference theory. *Proceedings of the National Academy of Sciences*, *114*(10), E1766–E1775.

Pearson, J. M., Watson, K. K., & Platt, M. L. (2014). Decision making: The neuroethological turn. *Neuron*, *82*(5), 950–965.

Pelé, M., Broihanne, M. H., Thierry, B., Call, J., & Dufour, V. (2014). To bet or not to bet? Decision-making under risk in non-human primates. *Journal of Risk and Uncertainty*, *49*(2), 141–166.

Pfanzagl, J., & Luce, R. D. (1959). Individual Choice Behavior. *Econometrica*.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–238.

Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, *22*(1), 134–142.

Prétôt, L., Bshary, R., & Brosnan, S. F. (2016). Comparing species decisions in a dichotomous choice task: adjusting task parameters improves performance in monkeys. *Animal Cognition*, *19*(4), 819–834.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, *3*(4), 323–343.

Quiggin, J. (1993). Rank-dependent Expected Utility - An Outline. In *Generalized Expected Utility Theory* (pp. 55–74). Springer Netherlands.

Raab, M., & Gigerenzer, G. (2015). The power of simplicity: A fast-and-frugal heuristics approach to performance science. *Frontiers in Psychology*, *6*(OCT), 1–6.

Rabin, M. (2000). Risk Aversion and Expected-utility Theory: A Calibration Theorem. *Econometrica*, *68*(5), 1281–1292.

Rabin, M., & Vayanos, D. (2010). The gambler's and hot-hand fallacies: Theory and applications. *Review of Economic Studies*.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. In *Trends in Cognitive Sciences*.

Rieger, M. O., & Wang, M. (2008). What Is Behind the Priority Heuristic? A Mathematical Analysis and Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, *115*(1), 274–280.

Rogers, J., & Gibbs, R. A. (2014). Comparative primate genomics: Emerging patterns of genome content and dynamics. *Nature Reviews Genetics*, *15*(5), 347–359.

Rudebeck, P. H., & Murray, E. A. (2014). *for Comparing and Contrasting Values*. 1–13.

Rustichini, A., Conen, K. E., Cai, X., & Padoa-Schioppa, C. (2017). Optimal coding and neuronal adaptation in economic decisions. *Nature Communications*, *8*(1).

Rustichini, A., & Padoa-Schioppa, C. (2015). A neuro-computational model of economic decisions. *Journal of Neurophysiology*, *114*(3), 1382–1398.

Samuelson, P. (1937). A Note on Measurement of Utility. In *Source: The Review of Economic Studies* (Vol. 4, Issue 2).

Santos, L. R., & Platt, M. L. (2013). Evolutionary Anthropological Insights into Neuroeconomics: What Non-Human Primates can Tell us About Human Decision-Making Strategies. *Neuroeconomics: Decision Making and the Brain: Second Edition*, 109–122.

Santos, L. R., & Rosati, A. G. (2015). The Evolutionary Roots of Human Decision Making. *Annual Review of Psychology*, *66*(1), 321–347.

Scheibehenne, B., Wilke, A., & Todd, P. M. (2011). Expectations of clumpy resources influence predictions of sequential events. *Evolution and Human Behavior*, *32*(5), 326–333.

Schoemaker, P. J. H. (1982). The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations. *Journal of Economic Literature*, *20*(2), 529–563.

Scholten, M., & Read, D. (2014). Prospect theory and the "forgotten" fourfold pattern of risk preferences. *Journal of Risk and Uncertainty*, *48*(1), 67–83.

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, *91*(6), 1402–1412.

Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews*, *95*(3), 853–951.

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32.

Schultz, W., Dayan, P., & Montague, R. P. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Seed, A., & Tomasello, M. (2010). Primate Cognition. *Topics in Cognitive Science*, *2*(3), 407–419.

Seo, H., Cai, X., Donahue, C. H., & Lee, D. (2014). Neural correlates of strategic reasoning during competitive games. *Science*, *346*(6207), 340–343.

Seo, H., & Lee, D. (2009). Behavioral and Neural Changes after Gains and Losses of Conditioned Reinforcers. *Journal of Neuroscience*, *29*(11), 3627–3641.

Shadlen, M. N., & Newsome, W. T. (2001). Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *Journal of Neurophysiology*, *86*(4), 1916–1936.

Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology*, *53*(1), 491–517.

Shi, Y., Cui, X., Yao, J., & Li, D. (2015). Dynamic Trading with Reference Point Adaptation and Loss Aversion. *Operations Research*, *63*(4), 789–806.

Simon, H. A. (1966). Theories of Decision-Making in Economics and Behavioural Science. In *Surveys of Economic Theory* (Vol. 49, Issue 3, pp. 1–28). Palgrave Macmillan UK.

Simon, H. A. (1990). Invariants of Human Behavior. *Annual Review of Psychology*, *41*(1), 1–20.

Smith, T. R., Beran, M. J., & Young, M. E. (2017). Gambling in rhesus macaques (Macaca mulatta): The effect of cues signaling risky choice outcomes. *Learning & Behavior*, *45*(3), 288–299.

Soltani, A., De Martino, B., & Camerer, C. F. (2012). A Range-Normalization Model of Context-Dependent Choice: A New Model and Evidence. *PLoS Computational Biology*, *8*(7), e1002607.

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty.

Nature Reviews Neuroscience, *20*(10), 635–644.

Stalmeier, P. F. M., & Bezembinder, T. G. G. (1999). The discrepancy between risky and riskless utilities: A matter of framing? *Medical Decision Making*, *19*(4), 435–447.

Starmer, C. (2000). Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *Journal of Economic Literature*, *38*(2), 332–382.

Starmer, C., & Sugden, R. (1989). Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty*.

Stauffer, W. R., Lak, A., Bossaerts, P., & Schultz, W. (2015). Economic Choices Reveal Probability Distortion in Macaque Monkeys. *Journal of Neuroscience*, *35*(7), 3146–3154.

Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error responses reflect marginal utility. *Current Biology*, *24*(21), 2491–2500.

Stevens, J. R. (2010). Rational Decision Making in Primates: The Bounded and the Ecological. In *Primate Neuroethology* (pp. 98–116). Oxford University Press.

Steward, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect Relativity: How Choice Options Influence Decision Under Risk. *Journal of Experimental Psychology: General*, *132*(1), 23–46.

Stott, H. P. (2006). Cumulative prospect theory’s functional menagerie. *Journal of Risk and Uncertainty*, *32*(2), 101–130.

Strait, C. E., Blanchard, T. C., & Hayden, B. Y. (2014). Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron*, *82*(6), 1357–1366.

Strier, K. B. (2011). *Primate behavioral ecology*. Prentice Hall.

Sugden, R. (1993). Review: Welfare, Resources, and Capabilities: A Review of Inequality Reexamined by Amartya Sen. *Journal of Economic Literature*, *31*(4), 1947–1962.

Sugden, R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory*.

Sugden, R., Kagel, J. H., Battalio, R. C., & Green, L. (2006). Economic Choice Theory: An Experimental Analysis of Animal Behaviour. *The Economic Journal*.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745–756.

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends in Cognitive Sciences*, *19*(1), 27–34.

Sundali, J., & Croson, R. (2006). Biases in casino betting: The hot hand and the gambler’s fallacy. In *Judgment and Decision Making* (Vol. 1, Issue 1).

Sutton, R., & Barto, A. (2018). Reinforcment Learning, Second Edition. In *MIT Press*. A Bradford Book.

Taswell, C. A., Costa, V. D., Murray, E. A., & Averbeck, B. B. (2018). Ventral striatum’s role in learning from gains and losses. *Proceedings of the National Academy of Sciences*, *115*(52), E12398–E12406.

Thaler, R. H. (2016). Behavioral Economics: Past, Present and Future. In *SSRN Electronic Journal* (Vol. 20, Issue 38).

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: improving decisions about health, wealth, and happiness*.

Timmins, R. J., Richardson, M., Chhangani, A. &, & Yongcheng, L. (2008). Macaca mulatta, Rhesus

Monkey. *The IUCN Red List of Threatened Species*, *8235*(T12554A3356486).

Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift Für Tierpsychologie*, *20*, 410–433.

Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2008). Neuronal Distortions of Reward Probability without Choice. *Journal of Neuroscience*, *28*(45), 11703–11711.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive Coding of Reward Value by Dopamine Neurons. *Science*, *307*(5715), 1642–1645.

Todd, P. M., & Gigerenzer, G. (2000). Précis of Simple heuristics that make us smart. In *Behavioral and Brain Sciences*.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. a. (2007). The neural basis of loss aversion in decision-making under risk-Supporting Material. *Science (New York, N.Y.)*, *315*(5811), 515–518.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, *398*(6729), 704–708.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*(2), 105–110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*, *59*(4), S251–S278.

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061.

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

Tymula, A. A., & Glimcher, P. W. (2016). Expected Subjective Value Theory (ESVT): A Representation of Decision Under Risk and Certainty. In *SSRN Electronic Journal*.

Uhlrich, P., Micheletta, J., Thierry, B., Dufour, V., & Pelé, M. (2010). Delay Maintenance in Tonkean Macaques (Macaca tonkeana) and Brown Capuchin Monkeys (Cebus apella). *International Journal of Primatology*, *32*(1), 149–166.

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)?: Research Article. *Psychological Science*, *20*(4), 473–479.

Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain. *Neuron*.

Volz, K. G., & Gigerenzer, G. (2014). The Brain Is not "As-If" – Taking Stock of the Neuroscientific Approach on Decision Making. In *Advanced Brain Neuroimaging Topics in Health and Disease - Methods and Applications* (Vol. 115, Issue 1, pp. 281–289). InTech.

Volz, K. G., Schooler, L. J., & von Cramon, D. Y. (2010). It just felt right: The neural correlates of the fluency heuristic. *Consciousness and Cognition*, *19*(3), 829–837.

von Neumann, J., & Morgenstern, O. (1944). Theory of games and economic behavior. In *Princeton University Press* (Issue 1).

von Neumann, J., Morgenstern, O., Kuhn, H. W., & Rubinstein, A. (2004). Theory of Games and Economic Behavior. In *John von Neumann and Oskar Morgenstern: Theory of games and economic behavior*.

Wainwright, M. J., Schwartz, O., & Simoncelli, E. P. (2001). *Natural Image Statistics and Divisive Normalization: Modeling Nonlinearities and Adaptation in Cortical Neurons*. MIT Press.

Wakker, P. P. (2010). Prospect theory: For risk and ambiguity. In *Prospect Theory: For Risk and Ambiguity*.

Watson, K. K., & Platt, M. L. (2008). Neuroethology of reward and decision making. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1511), 3825–3835.

Watzek, J., & Brosnan, S. F. (2018). (Ir)rational choices of humans, rhesus macaques, and capuchin monkeys in dynamic stochastic environments. *Cognition*, *178*(May), 109–117.

Webb, R. (2018). The (Neural) Dynamics of Stochastic Choice. *Management Science*, *65*(1), 230–255.

Webb, R., Glimcher, P. W., & Louie, K. (2014). *Rationalizing Context-Dependent Preferences : Divisive Normalization and Neurobiological Constraints on Choice*.

Weber, B. J., & Chapman, G. B. (2005). Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes*, *97*(1), 31–46.

Weber, E. U., & Johnson, E. J. (2008). *Mindful Judgment and Decision Making*.

Weber, M., & Camerer, C. F. (1987). Recent developments in modelling preferences under risk. *OR Spektrum*, *9*(3), 129–151.

Webster, M. A., Werner, J. S., & Field, D. J. (2005). Adaptation and the Phenomenology of Perception. In *Fitting the Mind to the World: Adaptation and After-Effects in High-Level Vision*.

Weiss-Cohen, L., Konstantinidis, E., Speekenbrink, M., & Harvey, N. (2018). Task complexity moderates the influence of descriptions in decisions from experience. *Cognition*, *170*, 209–227.

Wenner, L. M. (2015). Expected prices as reference points-Theory and experiments. *European Economic Review*, *75*, 60–79.

Wilke, A., & Barrett, H. C. (2009). The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evolution and Human Behavior*, *30*(3), 161–169.

Wilke, A., Hutchinson, J. M. C., Todd, P. M., & Czienskowski, U. (2009). Fishing for the right words: Decision rules for human foraging behavior in internal search tasks. *Cognitive Science*.

Wilke, A., Scheibehenne, B., Gaissmaier, W., McCanney, P., & Barrett, H. C. (2014). Illusionary pattern detection in habitual gamblers. *Evolution and Human Behavior*, *35*(4), 291–297.

Wilke, S. D., & Eurich, C. W. (2002). Representational accuracy of stochastic neural populations. *Neural Computation*.

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, *81*(2), 267–279.

Worthy, D. A., & Todd Maddox, W. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes. *Journal of Mathematical Psychology*, *59*(2), 41–49.

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

Wu, S.-W., & Glimcher, P. W. (2018). *The Emerging Standard Neurobiological Model of Decision Making* (S.-H. Chen, M. Kaboudan, & Y.-R. Du (eds.); Vol. 1, Issue November). Oxford University Press.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*(2), 140–176.

Xu, E. R., & Kralik, J. D. (2014). Risky business: Rhesus monkeys exhibit persistent preferences for risky options. *Frontiers in Psychology*, *5*(APR), 1–12.

Xu, J., & Harvey, N. (2014a). The hot hand fallacy and the gambler's fallacy: What are they and why do people believe in them? In *Problem Gambling: Cognition, Prevention and Treatment* (pp. 61–73). Palgrave Macmillan.

Xu, J., & Harvey, N. (2014b). Carry on winning: The gamblers' fallacy creates hot hand effects in online gambling. *Cognition*, *131*(2), 173–180.

Yaari, M. E. (2006). The Dual Theory of Choice under Risk. *Econometrica*.

Yamada, H., Louie, K., Tymula, A., & Glimcher, P. W. (2018). Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nature Communications*, *9*(1), 1–11.

Yamada, H., Tymula, A., Louie, K., & Glimcher, P. W. (2013). Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. *Proceedings of the National Academy of Sciences*, *110*(39), 15788–15793.

Yoo, S. B. M., & Hayden, B. Y. (2018). Economic Choice as an Untangling of Options into Actions. *Neuron*, *99*(3), 434–447.

Zakrzewski, A. C., Perdue, B. M., Beran, M. J., Church, B. A., & Smith, J. D. (2014). Cashing out: The decisional flexibility of uncertainty responses in rhesus macaques (Macaca mulatta) and humans (Homo sapiens). *Journal of Experimental Psychology. Animal Learning and Cognition*, *40*(4), 490–501.

Zheng, H., Wang, X. T., & Zhu, L. (2010). Framing effects: Behavioral dynamics and neural basis. *Neuropsychologia*, *48*(11), 3198–3204.

Zimmermann, J., Glimcher, P. W., & Louie, K. (2018). Multiple timescales of normalized value coding underlie adaptive choice behavior. *Nature Communications*, *9*(1), 1–11.