The description-experience gap: a challenge for the neuroeconomics of decision-making under uncertainty

Garcia Basile[1], Cerrotti Fabien[1], Palminteri Stefano[1]*

Abstract

The experimental investigation of decision-making in humans relies on two distinct types of paradigms, involving either description- or experience-based choices. In description-based paradigms decision variables (i.e., payoffs and probabilities) are explicitly communicated by mean of symbols. In experience-based paradigms decision variables are learnt from trial-by-trial feedback. In the decision-making literature 'description-experience gap' refers to the fact that different biases are observed in the two experimental paradigms. Remarkably, well-documented biases of description-based choices, such as under-weighting of rare events and loss aversion, do not apply to experience-based decisions. Here we argue that the description-experience gap represents a major challenge, not only to current decision theories, but also to the neuroeconomics research framework, which relies heavily on the translation of neurophysiological findings between human and non-human primate research. In fact, most non-human primate neurophysiological research relies on behavioural designs that share features of both description and experience-based choices. As a consequence, it is unclear whether the neural mechanisms discovered in non-human primate electrophysiology should be linked to description-based or experience-based decision-making processes. The picture is further complicated by additional methodological gaps between human and non-human primate neural research. After analysing these methodological challenges, we conclude proposing new lines of research to address them.

Keywords

neuroeconomics, description-experience gap, reinforcement learning, decision-making, macaque, risk

[1] *Laboratoire de Neurosciences Cognitives et Computationnelles, Ecole Normale Supérieure, Institut National de la Santé et Recherche Médicale, Université de Recherche Paris Sciences et Lettres, Paris, France.*

The neuroeconomic research program

The expected utility model was established as the standard normative model of decision-making under risk [1,2]. Integrating Bernoulli's intuition about the curvature of utility function and probability theories, Von Neumann and Morgenstern demonstrated that choices based on the expected utility (i.e., the product between utility of an outcome and its probability) satisfies four basic axioms of rationality (completeness, transitivity, continuity and independence). Historically, the neoclassical economics research program disregarded the study of the internal processes governing economic behaviours. Keynes' animal spirits [3] were considered unmeasurable, and economic theory was built on the assumption that the human mind as well the brain were ultimately black boxes. The 'as-if' hypothesis [4] illustrates this position by endorsing an instrumentalist epistemology: theory predictive power prevails on the realism of its initial assumptions. Accordingly, it was considered acceptable to rely on unrealistic assumptions regarding the unbounded cognitive capacities or perfect knowledge of economic agents, as far as the predictions were sufficiently accurate.

However, with the accumulation of behavioural evidence against the standard normative expected utility model, it soon appeared that it had to be profoundly amended to successfully account for decisions under risk [5,6]. Positive, descriptive, models of decision under risk that integrate insights from psychology, such as the notion of bounded rationality (i.e., humans display limited computational capacities), heuristics (taking computational shortcuts to make decisions) and biases (selectively overweighting some variable) were then proposed and formalized [7–9]. Among the descriptive theories of decision under risk and uncertainty, 'prospect theory' (PT) had a strong empirical ground and stood out [8,10]. PT postulates that expected utility is calculated relative to a reference point (the frame), an asymmetric treatment of gains and losses (loss aversion), as well as a subjective weighting of probabilities (probability distortion). PT proved successful to explain known paradoxes (such as the Allais's paradoxes) and new ones (e.g., the Asian disease paradox, as well as a certain

number of 'real life' irrational behaviours [11, 12]).

However, despite these successes, some aspects of the descriptive approach, in general, and PT, in particular, remained unsatisfactory. First, it remained difficult to ultimately arbitrate between competing descriptive theories solely based on behavioural data. For instance, alternative behavioural theories have been proposed (such as rank-dependent utility, regret and disappointment theories; see [13] for a review) that make overlapping predictions with PT, making them hard to disentangle. Second, while making accurate predictions, PT, and other descriptive theories, do not specify which are the actual cognitive operations and how they are implemented by the brain. In terms of the Marrian analysis of modeling, PT (as other descriptive theories) is situated at the computational level that specifies which is the goal of the agent (in this case: maximizing a subjective utility that includes reference point dependence, loss aversion and probability deformation), but is silent concerning the algorithmic (i.e., what are the operations involved in the manipulation of decision variables) and implementational levels (i.e., how these operations are physically embodied and realized) [14].

A couple of decades later the time was ripe for a group of scholars of diverse origins seeing in neuroscientific data the way to address the issues related to descriptive theories, developed by psychologists and behavioural economists. This was facilitated by the rapid development of non-invasive neuroimaging techniques in humans (most notably functional magnetic resonance imaging: fMRI [15-17]) and improvement of single unit electrophysiological recordings in monkeys [18,19]. The hope was (and still is) that, taking advantage of neuroscientific methods and concepts, neuroeconomics (as this raising field was named), would be able to address the epistemological issues of economic theories highlighted above. Concerning the adjudicating on competing theories (our first issue), by opening the brain "black box" functional neuroimaging studies would provide an additional crucial and observable measure - neural activity (as indexed by blood oxygen level dependent - BOLD - signal: an aggregate and indirect measure of neural electrical activity) to compare, ultimately falsify and refine behavioural models. We define this approach as the weak neuroeconomic agenda, as it does not involve rewriting economic descriptive theories [20–22]. Coming back to our example, while making similar behavioural predictions in respect of preferences under risk, different theories postulate different utility functions that can be searched in the brain [23–25]. Assuming one knows where to look for utility representation in the brain[1], it would be, in principle, possible to assess which model better predicts its activity (a sort of neural model comparison: see [29]). Beyond comparing different theories, neural activity could in principle help refining a theory by fixing some of its parame-

ters. For instance, in many circumstances PT is silent about how the reference point should be set [30]. Assuming one knows where to look for positive (gain) and negative (loss) utility representation in the brain, in some cases the reference point could be inferred comparing the profile of activity of the 'gains' and 'losses' areas[2] [25,33].

Concerning build new theories (second issue), accepting the fundamental epistemological tenet that (economic) decisions ultimately result from neural activity in the brain (which is a standard materialistic and monistic assumption with regards to the mind-body problem, see [34]), entails that neuroscientific methods should provide the conceptual and methodological tools necessary to develop new, neurobiologically grounded, neural models encompassing the algorithmic and implementational levels. In contrast with the previous approach, we define this approach as the strong neuroeconomic agenda, as it involves rewriting economic theories in neurobiological terms. By integrating biological constraints and cost functions, such neurobiologically grounded economic models, have the ambition of redefining why human decision-making presents certain biases from a biologically (not logically or statistically) normative perspective [35,36].

The methodological requirements of the two main neuroeconomics agenda are not quite the same. The weak neuroeconomic agenda can, in principle, be fulfilled by experiments relying on aggregate and indirect measures of the neural activity, such as the BOLD signal recorded by fMRI scanners in areas encoding subjective values. Furthermore, since goal is arbitrating between different behavioural theories of decision-making developed by psychologists and economists, the experiments belonging to this research agenda should be preferentially (if not exclusively) performed in humans.

On the other side, as neural models are, ultimately, models of which information is encoded in neurons and how neurons are connected (networks), the strong neuroeconomic agenda research program cannot be pursued only relying on fMRI neural signals[3]. In fact, BOLD signal, at its best of resolution, aggregates over thousands of neurons [37–39]. Furthermore, it is still unclear to which extent it reflects presynaptic or postsynaptic activity (probably a mixture of both) [39,40]. Such neural models should eventually be validated based on recording of single cell activities, which is, for obvious ethical reasons, nearly impossible in humans[4]. This is why neuroeconomics research, from its very inception, strongly relies on electro-

---

[1]Subjective utility (or subjective value) representation seems to be distributed across a network of areas that include the ventral and the dorsal prefrontal cortices (both medial and lateral part), posterior cingulate cortex, the striatum, the insula, the amygdala and the hippocampus [26–28].
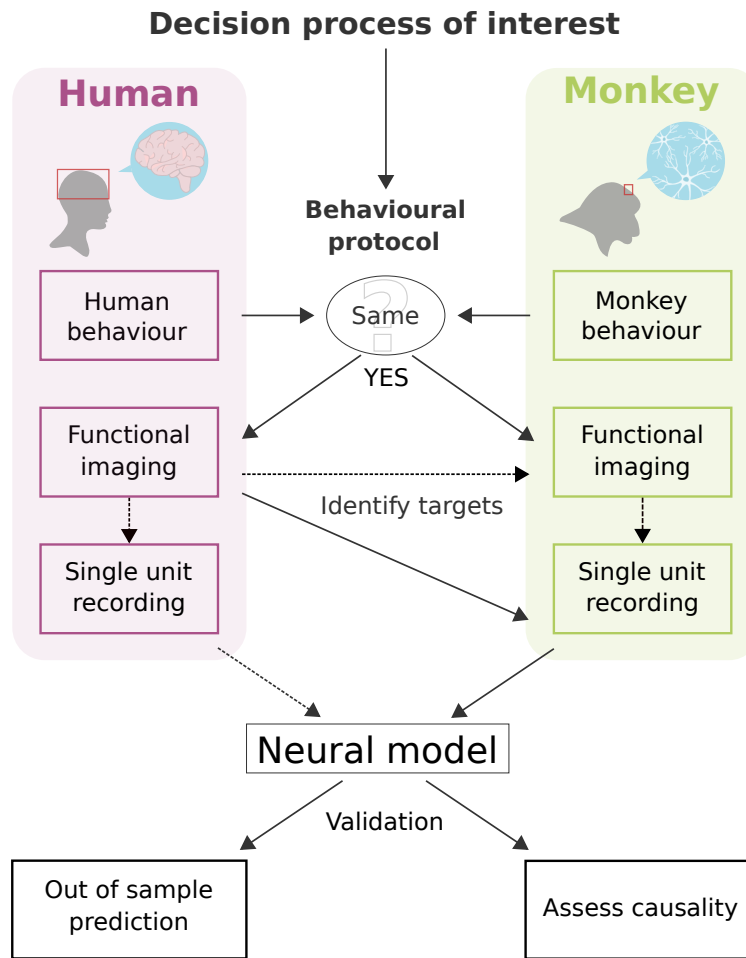
[2]It is indeed the case that brain systems encoding positive and negative values are, at least partially, dissociable. Losses are generally encoded by the insula, the amygdala and the dorsal prefrontal cortex, while gains are generally encoding in the ventral prefrontal and striatal [31,32].

[3]Other non-invasive imaging techniques, such as magneto- and electroencephalography present no advantage over fMRI when it comes to infer single unit activity. They present better temporal resolution traded off against a worst spatial resolution.

[4]There are few exceptions of single unit recordings in humans, obtained from neurologic patients undergoing brain surgery, while informative, these data are limited in the fact the neuro-anatomical targets cannot be chosen freely and that neurological diseased brain may not represent the general population [44].

Figure 1. Prototypical workflow combining human (purple) and monkey (green) data to pursue the strong neuroeconomic agenda. Dotted lines designate optional steps.

physiological research on animal models, which have been employed in the study of neural mechanisms and cognition for almost 80 years [41]. Monkeys (especially rhesus monkey: macaca mulatta), are particularly popular models, because they present a wide behavioural repertoire and high degree of neuro-anatomical homology with humans, especially concerning the prefrontal cortices that underpin decision-making [42].

In **Figure 1** we represent how a prototypical workflow should look like to combine human and monkey data to deliver a neural model of decision-making. Of note, we describe it from an abstract perspective of theory-building, but in reality, its different steps can occur simultaneously (or in reverse order), and in very distant laboratories. Once identified as a behavioural process of interest (e.g., decision-making under uncertainty) a behavioural protocol is designed (typically, a series of choice problems involving different amounts of rewards and probabilities) and administered to both humans and monkeys. If the behaviour is comparable across species (meaning that the monkey represent a valid experimental model of hu-

man behaviour[5]), functional imaging in humans can then be deployed to identify neural targets encoding macroscopic variables (e.g., probabilities, outcomes) that are later used to guide the selection of the areas where single units will be recorded in monkeys. As desirable intermediate step, to reinforce the functional correspondence between human and monkey brain activations would be to deploy fMRI also in monkey [43]. Similarly, in some neurologic and psychiatric diseases, intra-cranial neural activity can also be recorded in humans [44]. Finally, all these data can then be combined together to propose and validate a neurobiologically plausible model of the behavioural process of interest. Thereafter, the proposed model should be validated using lesions and assessing its generalization . Methods such as trans-cranial magnetic stimulation and brain lesions can be used to test the alleged

---

[5]Of course, there is a lot of information to be gained also in the case where humans and monkeys do not display the same decisions and biases. Such differences currently represent a strong area of research in comparative psychology and ethology [71]. However, the (not so implicit) assumption of the vast majority of research in neuro-economics is that monkeys are valid experimental models for human cognition, and they are not investigated for comparative reasons.
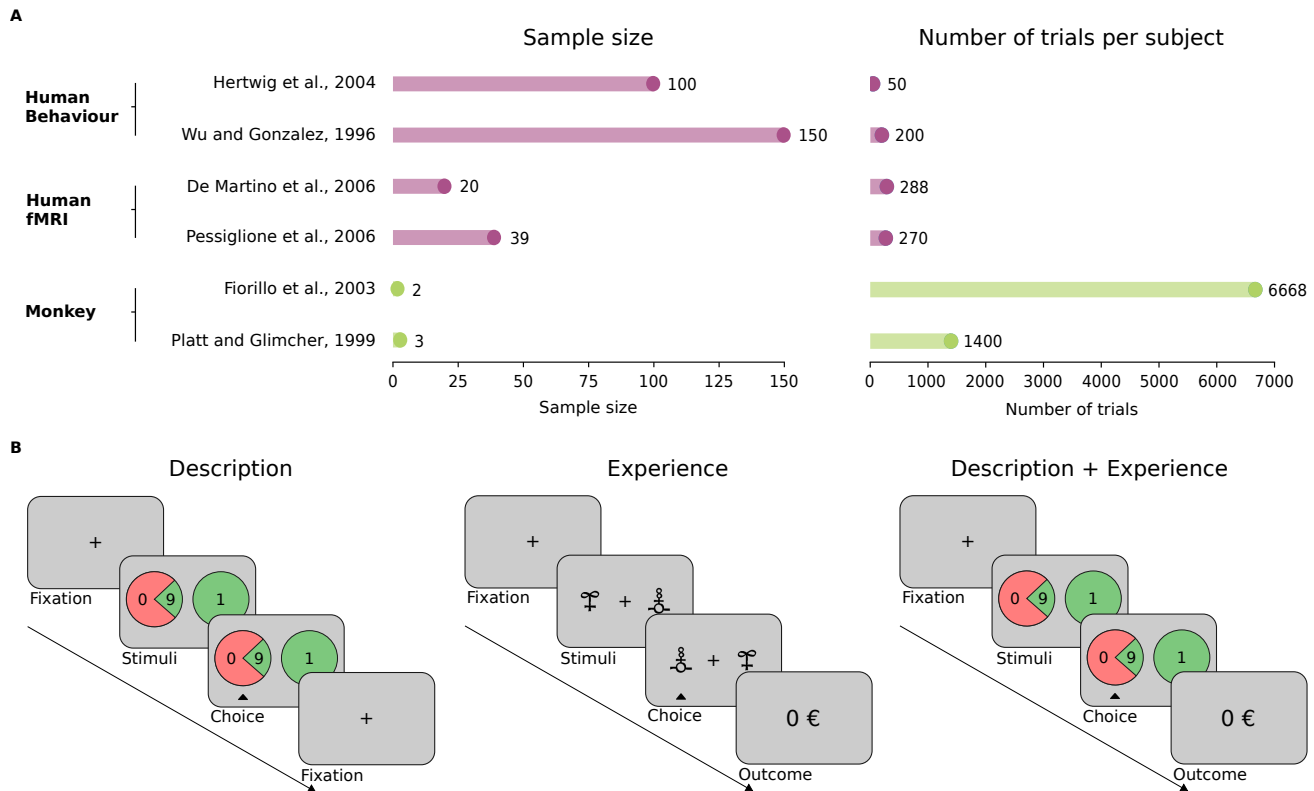
Figure 2. Methodological differences between description, experience and description-experience studies. **A**, Sample size and number of trials listed in two electrophysiological studies [51,52], two human fMRI studies [25,55] and two human behavioural studies [53,54]. **B**, Schematic representation of the different steps forming a trial in these different paradigms. In pure description, two options are presented with explicit probability distribution of outcomes, while no feedback is given (outcome of the decision). In pure experience, two abstract cues are presented without any a priori information, but subjects are provided with the outcome of their decisions. In the 'description plus experience' paradigm, explicit representations of probabilities and outcomes are provided, as well as the outcome obtained subsequently to the choice.

causal relationship between neural correlates and behavioural processes [45–47]. The model's ability to generalize can be assessed by generating predictions in tasks involving different decision-problems and behavioural processes (out-of-sample validation). A crucial step in this workflow is checking that humans and monkeys display the same behavioural processes and biases as a result of a true homology. This is something notoriously tricky to assess, because several, to some extent unavoidable, methodological differences exist between human and non-human primate research.

The foundational experimental paradigm of behavioural decision-making research consists in making choices between 'lotteries' or 'gambles', i.e., options associated with known or unknown probabilities of obtaining different outcomes [2,5]. According to the gambling metaphor of individual choice [48], lotteries are believed to be prototypical of real-life decisions [49]. Outcomes and their probabilities are described to participants, who often (especially in the first generation of behavioural economics studies) make only one or very few choices, without being informed about the result of their choices (sometime to purposely avoid learning processes to

impact decision-making [50]). On the other side, monkey electrophysiological research adopts very different methodological standards. For various reasons (including ethical ones), monkey studies are limited in terms of sample size, and consequently the number of observations per subject is greatly increased in order to improve statistical power and reduce noise in measurements. In fact, behavioural tasks in monkeys display a greater number of trials per subject, collected on a sample size of often less than five subjects (e.g., [51,52]). Both parameters (sample size and number of trials) are roughly a couple of orders of magnitude different compared to what is common practice in behavioural economics (e.g., [53,54]) (**Figure 2A**).

Strikingly, fMRI studies of decision-making present experimental parameters somehow in-between those used in monkeys and Human studies: they usually involve hundreds of trials and also sample sizes of about 20-40 subjects (see two notable examples in neuroeconomics: [25,55]). Assuming that decision-making possesses ergodicity (i.e., the behaviour averaged across trials is the same as the behaviour averaged across subjects), different ratio trial/participants per se should
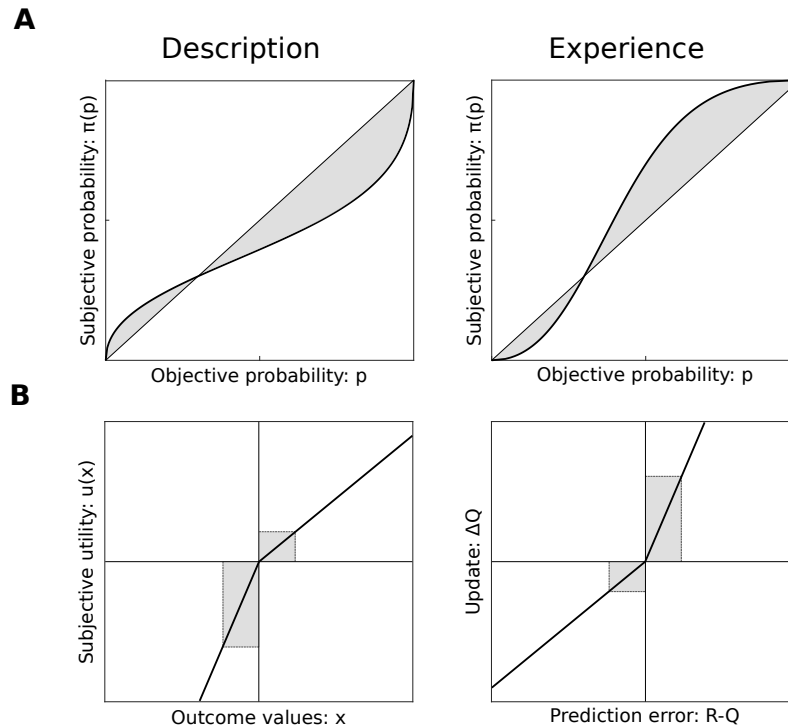
A

Description | Experience

B

Figure 3. A, Illustration of the non-linear transformation of probabilities in description (left panel) and experience (right panel). In the description domain, subjective probability is reflected by a probability weighting function (here denoted π) following an inverse S-Shape (i.e. low probabilities are overweighted while high probabilities are underweighted). This tendency is reversed when it comes to the experience domain, where the curve follows an S-shape. **B**, Illustration of the classical linear utility function in the description domain (left panel) and the value function for the experience domain (right panel). In description, the utility curve displays a steeper slope for losses than gains. In experience, an opposite phenomenon is frequently observed. The sign of the prediction error (i.e. the difference between the obtained reward R and the associative value Q) affects learning.

not present a big challenge to compare results from human and monkey studies (but note that ergodicity does not seem to be granted for psychological processes, see [56]). However, in addition to these quantitative differences, in monkey studies, an outcome (usually a primary reward) is provided on a trial-by-trial basis. This is because a monkey would simply stop doing the experiment in the absence of extrinsic motivation. Thus, in virtual all cases monkey experiments include a reinforcement learning component, where past actions are associated with past outcomes. This is true even when the paradigm involves establishing a symbolic system to communicate outcomes and probabilities. In fact, in absence of a shared language or semantic system to communicate, monkeys are compelled to learn any representational system by trial-and-error from feedback.

In the present article we argue that the above mentioned differences do not only present a technical issue, but also a major epistemological challenge for the (strong) neuroeconomic agenda. We detail below why.

The description-experience gap

As mentioned before, foundational contributions to decision behavioural science were made through the use of explicitly described gambles. Several representations have been used to convey outcomes values and probabilities, including textual and numerical descriptions (e.g. [5,8,53]), later replaced by visual cues such as pie-charts (e.g. [25,57]). In these paradigms the information pertaining the decision-relevant is processed by verbal and mental calculation systems and rely upon some degree of semantic knowledge to decode the meaning to the symbols used. In addition to that, decision problems were usually presented only once and, in case multiple decision problems were used, the final outcome (i.e., the realisation of the lottery) was usually not displayed on a trial-by-trial basis (**Figure 2B**). However, relatively few situations in real life match the characteristics of the pure description-based paradigms, namely a complete and explicit information about outcome values probabilities. In fact, in many circumstances, it seems rather prudent to assume that information about outcome values and probabilities are shaped by past encounters of the same decision problem. Experimentally, this configuration is often translated into multi-armed bandit problems (starting with Thompson [58], but see [59] for a review), where the decision-maker faces abstract cues of unknown value and has to figure by trial-and-error the value of the options. Computationally, behaviour in multi-armed bandit problems is

generally well-captured by associative or reinforcement learning processes [60]. In the early 2000s a line of enquiry arose where researcher translated the typical decision problems used in behavioural economics (i.e., involving choices between a safe and a risky prospect in the gain and loss domain[6]) into experience-based paradigms [54,62,63] (**Figure 2B**). Systematic comparisons between these two decision-making modes revealed the existence of robust description-experience gaps regarding risk preferences in humans [64–66] . More precisely, probability weighting functions eventually show opposite deformations when comparing description-based and experience-based choices (**Figure 3**, **Box 1**). In particular, most of the tenets of prospect theory do not seem to hold in experience-based choices [8]. While traditionally, in the description domain, the occurrence of rare events is overestimated (possibility effect) and the occurrence of frequent events underestimated, experience-based decision tends to show the opposite biases: an effect that is only partially explained by incomplete sampling [54,62,63].

In description-based choices, a behavioural hallmark of loss aversion (overweighting of negative outcomes) is the reflection effect, where subjects are risk-averse in the gain domain and risk-seeker in the loss domain. The opposite pattern has been repeatedly found in the experience-based decisions [66]. This observation may be explained by biases in the learning process, such as remembering preferentially extreme outcomes or integrating preferentially better-than-expected outcomes [67,68]. Finally, a smaller subset of studies investigated a hybrid situation where decision problems are fully described, choices are repeated and followed by a trial-by-trial basis feedback. These 'description plus experience' paradigms showed that prospect-theory compatible probability distortions are initially present, but corrected by the presence of feedback [69,70]. To summarize, the whole spectrum of decision-making under uncertainty in humans is far from being fully captured by PT's loss aversion and subjective probability deformation. Specifically, different descriptive models seem to apply as a function of how outcome and probability information is conveyed. In the remaining of the paper, we illustrate why we believe that this feature seriously challenges leveraging on neural and behavioural data in monkeys to build a neural model of decision-making under uncertainty.

Decision under risk in monkeys

In the present section we try to address the question of whether monkeys are a good experimental model for human decision-making under uncertainty. We will focus this survey on rhesus monkey (macaca mulatta) results because most electrophysiological studies are performed in this species (but see [71]

for a more detailed review including other primates). Asking whether monkeys are a good experimental model translates into asking whether in the laboratory setting their behaviour displays the distinctive features and biases observed in humans. We stress again that the comparison is complicated by the fact that pure description-based paradigms cannot exist in monkey studies because of the lack of language. In fact, whenever in monkey studies, outcomes and probabilities are conveyed via a symbolic system, it is still learned and maintained by trial-by-trial outcomes (i.e., a situation similar to the 'description plus experience' paradigm, described above). In such 'pseudo' description-based paradigm, monkeys are trained to associate continuous variations in one visual feature (e.g., colour and size) to continuous variations of a decision variable (e.g., outcomes or probabilities). The comparison is further complicated by the fact that only few studies formalize risk preferences in terms of model parameters (such as probability distortion, loss aversion or learning rates) and data reporting is often limited to behavioural measures.

The general picture (**Table 1**) emerged from "pseudo" description-based paradigms in monkeys (i.e., studies relying on learned symbolic systems to communicate values) is, at best, mixed. Prospect theory has been explicitly tested in paradigms using visual cues carrying symbolic information similar to those presented to humans (e.g. pie-charts). Only few studies show results in conformity with the pattern of description-based decisions observed in humans. Risk-aversion, suggestive of marginally decreasing utility in the gain domain, has been rarely reported [72]. Nioche et al. [73] is the sole study confirming all PT features: marginally decreasing utility (risk aversion in the gain domain), loss aversion (risk seeking in the loss domain) and subjective probability weighting consistent with overestimation of rare events. Probability weighting function consistent with standard PT has been reported by other studies, but the same studies also reported increasing marginal utility and risk seeking in the gain domain, which is not typically observed in description-based decisions in humans [74,75]. Many others pseudo description-based experiments also reported risk-seeking attitudes and/or marginally increasing utility in gains [76–79]. In addition, although the traditional inverse probability weighting function has sometimes been observed [73,74], variation of experimental design features (such as randomly mixing gambles instead of repeating the same gambles sequentially) can reverse the direction of the probability weighting function [80].

Regarding "pure" experience-based studies in monkeys (i.e., involving no symbolic system to communicate values), the picture is somehow clearer. Indeed, rhesus macaques exhibit robust risk-seeking behaviour in the gain domain [81–90]. Risk-seeking attitudes have also been reported in the loss domain [91]. Risk-seeking behaviour in experience-based studies can be computationally explained by an increased sensitivity to positive (compared to negative) prediction errors ('positivity' bias) as the one generally documented in humans (**Box 1**) [67,92,93]. This hypothesis is corroborated by studies

---

[6]In the human reinforcement learning literature, the most frequently used paradigms involve options that possess, at a given trial, different expected values but overall similar risk level [55,61]. As a result the human reinforcement learning literature is more concerned about measure of objective performance rather than subjective preference.

Study	Sample size	Modality	Reward	Risk attitude in gains	Risk attitude in losses	Probability distortion	Loss aversion
McCoy & Platt (2005)	2	E	Juice	Seek	N/A	N/A	N/A
Hayden & Platt (2007)	2	E	Juice	Seek	N/A	N/A	N/A
Hayden et al. (2008)	5	E	Juice	Seek	N/A	N/A	N/A
Long (2009)	3	E	Juice	Seek	N/A	N/A	N/A
Watson (2009)	8	E	Juice	Seek	N/A	N/A	N/A
O'Neill & Schultz (2010)	2	E	Juice	Seek	N/A	N/A	N/A
Heilbronner et al. (2011)	3	E	Juice	Seek	N/A	N/A	N/A
Kim et al. (2012)	2	E	Juice	Seek	N/A	N/A	N/A
Heilbronner & Hayden (2013)	2	E	Juice	Seek	N/A	N/A	N/A
Xu & Kralik (2014)	2	E	Juice	Seek	N/A	N/A	N/A
Smith et al. (2017)	7	E	Juice	Seek	Seek	N/A	N/A
Hayden et al. (2010)	4	D	Juice	Seek	N/A	N/A	N/A
So & Stuphorn (2010)	2	D	Juice	Seek	N/A	N/A	N/A
Yamada et al. (2013)	2	D	Juice	Avoid	N/A	N/A	N/A
Raghuraman & Padoa-Schioppa (2014)	2	D	Juice	Seek	N/A	N/A	N/A
Staufer et al. (2015)	2	D	Juice	Seek	Seek	Inverse S-Shape	N/A
Farashahi et al. (2018), experiment 1	3	D	Juice	Seek	N/A	Linear Shape	N/A
Farashahi et al. (2018), experiment 2	3	D	Token	Seek	Seek	S-Shape	N/A
Chen, Stuphorn (2018)	2	D	Juice	Seek	Seek	Inverse S-Shape	N/A
Nioche et al. (2019)	2	D	Juice	Avoid	Seek	Inverse S-Shape	Yes
Ferrani-Toniolo et al. (2019), experiment 1	2	D	Juice	N/A	N/A	Inverse S-Shape	N/A
Ferrani-Toniolo et al. (2019), experiment 2	2	D	Juice	N/A	N/A	S-Shape	N/A
Eisenreich et al. (2019)	3	D	Juice	Seek	Seek	N/A	N/A

Table 1. Studies investigating risk attitudes in rhesus monkey. The modality 'E' denotes experience paradigms (i.e. without explicit representation of outcomes and probabilities). The modality 'D' denotes description paradigms (i.e. involving explicit representation of outcomes and probabilities). Note that in monkeys this implies a 'description plus experience' set-up. We mention the nature of the reward, which is most of the time a juice (fruit juice or water). Risk attitudes can take two values, either 'seek', either 'avoid'. When probability distortions are studied, we mention the direction of the curve (either an S-Shape, either an inverse S-Shape). When loss aversion is explicitly studied and replicated, we indicate it.

demonstrating a stronger impact of past positive outcome in choices using either model free and model-based measures [82,85,94].

Finally, it can be argued that if monkeys are a good model for human decision-making under uncertainty, they should display a description-experience gap. To our knowledge, so far only one study started explicitly tackling this issue [95]. Monkeys were asked to make repeated choices between safe, and risk options, whose outcome probability was either learned by experience or described by the ratio between colors on a rectangle. Replicating previous findings in monkeys, and in discordance with the standard result in humans, Heilbronner and Hayden found that monkeys were risk-seekers in the description domain. However, consistent with the gap observed in humans, they also found that risk seeking behaviour was higher for experience-based cues.

To summarize, the literature seems to suggest that monkeys' decision-making for experience-based choice is quite consistent to what is observed in humans in terms of risk-preference. This is consistent with a large body of literature showing that the neural substrates of reinforcement learning are largely preserved in the two species [96,97]. Risk seeking in this context may be driven by higher learning rate from positive compared to negative prediction errors, which is essentially a computational reinforcement learning translation of the 'hot hand' fallacy [98,99]. The situation is much less promising concerning description-based decisions, as preferences compatible with prospect theory are rarely observed. This can be due to the fact that pseudo description-based design in monkeys resembles the 'description plus experience' set-up in humans, where PT-like deformation are blunted or even disappear, as if description-based and experience-based biases reciprocally cancel themselves [69,70]. As a result, it remains unclear to what extent description-based processes can be captured by the monkey model.
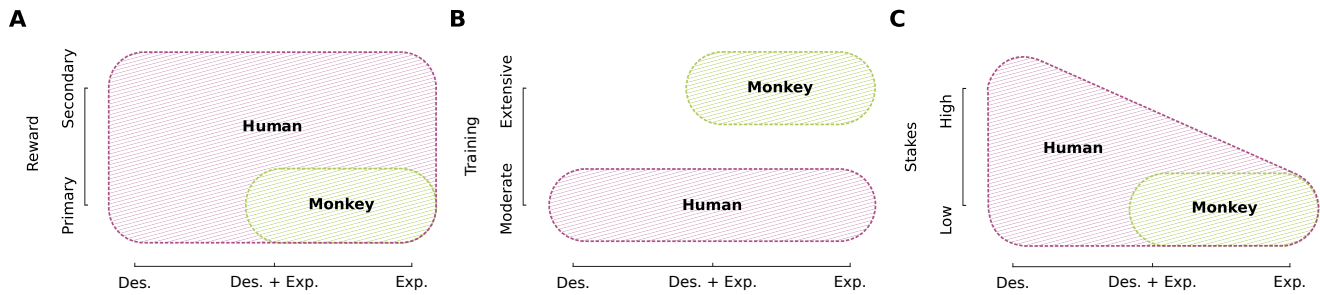
Figure 4. The figure illustrates how human (purple) and monkey (green) experimental settings map into a four-dimensional space, whose axes are: the way value information is provided (from description to experience); the nature of the reward (from primary to secondary; **A**), the amount of training (from moderate to extreme; **B**) and the amount of the stakes (from low to high; **C**).

The impact of other experimental differences

Experimental results concerning decision-making under uncertainty in monkeys do not seem to straightforwardly comply with the predictions of prospect theory. Overall it seems that monkeys' behaviour is better accounted for as an experience-based decision process, which is consistent with the fact that pure description-based paradigms are not possible and monkey experiments always involve trial-by-trial feedback. The systematic presence of trial-by-trial feedback is not the only recurrent methodological difference between the monkey and human studies (**Figure 2** and **Figure 4**).

First, monkey studies essentially rely on primary rewards (mainly water or fruit juice), while human studies are realized mainly with secondary rewards (sometimes hypothetical ones) and primary reinforcers are only occasionally used [100,101]. Preliminary evidence from a study comparing risk propensity for different kinds of rewards in humans (money versus sport beverage) and monkeys showed similar patterns, thus suggesting that in the good experimental condition risk preferences in both species could converge [102]. Furthermore, while the neural correlates of different kinds of rewards converge in the ventral prefrontal and striatal systems (principle of the common currency; [103]) they also have specific correlates, which may contribute to the different neural mechanisms and result in distinct, reward-specific, risk preferences [100]. On the other side, a proxy for secondary reward can be found in monkey paradigms that involve collecting (virtual) tokens to be later exchanged for a primary reward. Unlike pure primary reward tasks, where losses cannot be implemented (it is impossible to take some fruit juice away from the stomach of a monkey), tokens have the advantage of making possible subtracting previously acquired rewards from the animal, thus inducing 'losses' in the same manner as in human. However, a recent study using tokens, showed risk-seeking attitudes comparable to that observed using primary reward [79]. Furthermore, when tokens are used, they are almost immediately changed against primary reward, making them not really comparable to money, whose value is much more

permanent. Taken together the available evidence suggests that the primary / secondary reward dichotomy per se does not explain the fact that human description-based biases are hardly observed in monkeys.

Second, in addition to the difference in the nature of the reward; description-based paradigms in humans and paradigms in monkeys often present a systematic difference in the amount of the reward (**Figure 4**). Indeed, most of the original studies about prospect theory used hypothetical gambles of hundreds of dollars and the same biases have been replicated using real stakes of about a month salary [104]. On the other side monkey studies use very small amounts of rewards (mere drops of liquids). It has been argued that the part that the description-experience gap may simply derive from this difference in stake and not present fundamental differences in the process that generates preferences [89]. This would be consistent with Markowitz utility function which suppose risk-seeking for small stakes ('peanuts effect') before converting to risk aversion of higher stakes [105] and is supported by the finding that increasing the relative amount of reward (by reducing its frequency) decreases risk-seeking down to risk neutrality in monkeys [89,105]. However, risk-aversion in the gain domain (and a reverse pattern in the loss domain: the reflection effect) has also been observed with small stakes in description-based decisions in humans [66]. Thus, available evidence suggests that differences in the size of the stake cannot fully explain the fact that human description-based preferences are hardly observed in monkeys.

Finally, another notable difference between human and monkey experiments is represented by the amount and the type of training required to perform the task (See **Figure 2** and **Figure 4**). In human experiments, task training rarely takes more than a few minutes (in some extreme cases of description-based paradigms, there is virtually no training, subjects are just asked their preferences). On the other side, monkey experiments requirement extensive training, in general spanning through several months (usually training takes longer than the experiment itself). It can be therefore argued that their behaviour becomes to some extent habitual or automatized: a cognitive state that contrasts dramatically with

the declarative and deliberative stance of description-based choices taken by humans [106]. In addition to that, training in monkeys (and other animals) often involves simplified versions of the task (often deterministic contingencies), which may reinforce specific risk preferences. Although the role of extended (several days, weeks) training and the resulting behavioural automation (or habituation) in risk preferences is unclear, it may contribute to the fact that human description-based biases are hardly observed in monkeys.

Conclusions and perspectives

Our review suggests that the rhesus monkey is a partial model of human decision-making under uncertainty. Risk preferences in monkeys are generally better explained as experience-based processes. Accordingly, monkeys proved to be a very good model of human reinforcement learning processes, providing crucial insights into its neural implementation (the dopamine prediction error hypothesis: [55,61,107]). The situation is less clear concerning description-based choices. In paradigms using explicit symbolic information about decision variables, only rarely monkeys displayed risk preferences compatible with human results. Deciding by description implies a symbolic system of communication. While in humans this system pre-exists (language), in monkeys it has to be learnt by trial-and-error, thus irremediably confounding description and experience. In addition to differences in the way value information is conveyed (experience- or description-based), other methodological factors (training, reward type and stakes) further drive away the experimental set-ups of the two species. This situation is problematic as building a neural model of decision-making under uncertainty should integrate human (fMRI) and monkey (single unit) neurophysiological data, while explaining risk preferences in a wide range of situations that span from pure description-based choices to pure experience-based choices.

We propose further lines of research that could eventually help filling these gaps and ultimately fulfilling the strong neuroeconomic agenda. On the human side, the description-experience gap has been extensively studied at the behavioural level, but surprisingly neglected at the neural level. A notable exception [108], found different neural representations for description- and experience-oriented decisions. Furthering this line of enquiry would prove useful to redefine the target areas where to look specifically for description-based processes in monkey electrophysiological studies.

With the development of online testing techniques, it is becoming easier to implement extended massive training in humans [109]. Translated in the field of decision-making under risk, these experiments would provide crucial insights into the impact of extensive training in risk preferences. While, description-based studies in monkeys, require learning ex novo a symbolic system, in humans the meaning of pie charts is provided by language. It would be interesting to put humans in situations, where they have to figure out by trial-and-error

the code linking continuous visual features to decision variables.

In general, all the efforts aimed at increasing the methodological overlap between human and monkey studies will provide further insights into what are the behavioural processes shared across the two species. Popularizing fMRI experiments in monkeys would help confirm the neuro-anatomical targets and increase the focus on shared neural systems. The token paradigm (conceptually closer to notion of secondary reward) offers the possibility to implement losses in monkeys, hence facilitating the cross-species study of loss aversion. Finally, on the monkey side, PT has been sporadically replicated. It will be important to understand and formalize the experimental factors (in terms of stimuli, training and reward type; see Table 1) that predict whether PT-like behaviour will be observed in a monkey experiment [89]. Determining under which experimental conditions prospect theory is replicated in monkeys, will imply a deeper understanding of the cognitive mechanisms underlying decision-making under uncertainty.

References

[1] Samuelson PA. A Note on the Pure Theory of Consumer's Behaviour. Economica. 1938;5(17):61–71.

[2] Von Neumann J, Morgenstern O. Theory of games and economic behavior. Princeton, NJ, US: Princeton University Press; 1944 xviii, 625 p. (Theory of games and economic behavior).

[3] Keynes JM. The General Theory of Employment, Interest, and Money. Springer; 1936. 430 p.

[4] Friedman M. Essays in Positive Economics. University of Chicago Press; 1953. 336 p.

[5] Allais M. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. Econometrica. 1953;21(4):503–46.

[6] Ellsberg D. Risk, Ambiguity, and the Savage Axioms. Q J Econ. 1961;75(4):643–69.

[7] Simon HA. A Behavioral Model of Rational Choice. Q J Econ. 1955 Feb 1;69(1):99–118.

[8] Tversky A, Kahneman D. Prospect theory: An analysis of decision under risk. Econometrica. 1979;47(2):263–91.

[9] Kahneman D, Slovic SP, Slovic P, Tversky A. Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press; 1982. 574 p.

[10] Tversky A, Kahneman D. Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertain. 1992 Oct 1;5(4):297–323.

[11] Kahneman D, Knetsch JL, Thaler RH. Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. J Econ Perspect. 1991 Mar;5(1):193–206.

[12] Camerer CF. Prospect theory in the wild: Evidence from the field. 1998;

[13] Vlaev I, Chater N, Stewart N, Brown GDA. Does the brain calculate value? Trends Cogn Sci. 2011 Nov;15(11):546–54.

[14] Marr D, Poggio T. From understanding computation to

understanding neural circuitry. 1976;

[15] Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RS. Spatial registration and normalization of images. Hum Brain Mapp. 1995;3(3):165–89.

[16] Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp. 1994;2(4):189–210.

[17] Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited–again. NeuroImage. 1995 Sep;2(3):173–81.

[18] Nordhausen CT, Maynard EM, Normann RA. Single unit recording capabilities of a 100 microelectrode array. Brain Res. 1996;726(1–2):129–40.

[19] Baker JT, Patel GH, Corbetta M, Snyder LH. Distribution of activity across the monkey cerebral cortical surface, thalamus and midbrain during rapid, visually guided saccades. Cereb Cortex. 2006;16(4):447–59.

[20] Camerer CF, Loewenstein G, Prelec D. Neuroeconomics: Why economics needs brains. Scand J Econ. 2004;106(3):555–79.

[21] Camerer C, Loewenstein G, Prelec D. Neuroeconomics: How neuroscience can inform economics. J Econ Lit. 2005;43(1):9–64.

[22] Rustichini A. Neuroeconomics: what have we found, and what should we search for. Curr Opin Neurobiol. 2009 Dec;19(6):672–7.

[23] Chua HF, Gonzalez R, Taylor SF, Welsh RC, Liberzon I. Decision-related loss: regret and disappointment. Neuroimage. 2009;47(4):2031–40.

[24] Coricelli G, Critchley HD, Joffily M, O'Doherty JP, Sirigu A, Dolan RJ. Regret and its avoidance: a neuroimaging study of choice behavior. Nat Neurosci. 2005;8(9):1255–62.

[25] De Martino B, Kumaran D, Seymour B, Dolan RJ. Frames, biases, and rational decision-making in the human brain. Science. 2006;313(5787):684–7.

[26] Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M. An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. Neuron. 2009 Nov 12;64(3):431–9.

[27] Bartra O, McGuire JT, Kable JW. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage. 2013;76:412–27.

[28] Clithero JA, Rangel A. Informatic parcellation of the network involved in the computation of subjective value. Soc Cogn Affect Neurosci. 2014;9(9):1289–302.

[29] Palminteri S, Khamassi M, Joffily M, Coricelli G. Contextual modulation of value signals in reward and punishment learning. Nat Commun. 2015;6(1):1–14.

[30] Baillon A, Bleichrodt H, Spinu V. Searching for the Reference Point. Manag Sci. 2020 Jan;66(1):93–112.

[31] Palminteri S, Pessiglione M. Opponent brain systems for reward and punishment learning: causal evidence from drug and lesion studies in humans. In: Decision Neuroscience. Elsevier; 2017. p. 291–303.

[32] Pessiglione M, Delgado MR. The good, the bad and the brain: neural correlates of appetitive and aversive values underlying decision making. Curr Opin Behav Sci. 2015;5:78–84.

[33] Tom SM, Fox CR, Trepel C, Poldrack RA. The Neural Basis of Loss Aversion in Decision-Making Under Risk. Science. 2007 Jan 26;315(5811):515–8.

[34] Bunge M. The mind–body problem: A psychobiological approach. Elsevier; 2014.

[35] Glimcher PW. Foundations of neuroeconomic analysis. OUP USA; 2011.

[36] Padoa-Schioppa C, Rustichini A. Rational attention and adaptive coding: a puzzle and a solution. Am Econ Rev. 2014;104(5):507–13.

[37] Goense J, Bohraus Y, Logothetis NK. fMRI at High Spatial Resolution: Implications for BOLD-Models. Front Comput Neurosci [Internet]. 2016 Jun 28;10.

[38] Iranpour J, Morrot G, Claise B, Jean B, Bonny J-M. Using High Spatial Resolution to Improve BOLD fMRI Detection at 3T. Yap P-T, editor. PLOS ONE. 2015 Nov 9;10(11):e0141358.

[39] Kayser C, Logothetis NK. The Electrophysiological Background of the fMRI Signal. In: fMRI. Springer; 2010. p. 23–33.

[40] Logothetis NK. What we can do and what we cannot do with fMRI. Nature. 2008;453(7197):869–78.

[41] Jacobsen CF. Studies of cerebral function in primates. I. The functions of the frontal association areas in monkeys. Comp Psychol Monogr. 1936;

[42] Kennerley SW, Walton ME. Decision making and reward in frontal cortex: complementary evidence from neurophysiological and neuropsychological studies. Behav Neurosci. 2011;125(3):297.

[43] Fouragnan EF, Chau BKH, Folloni D, Kolling N, Verhagen L, Klein-Flügge M, et al. The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. Nat Neurosci. 2019 May;22(5):797–808.

[44] Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, et al. Human Substantia Nigra Neurons Encode Unexpected Financial Rewards. Science. 2009 Mar 13;323(5920):1496–9.

[45] Ruff CC, Driver J, Bestmann S. Combining TMS and fMRI: From 'virtual lesions' to functional-network accounts of cognition. Cortex. 2009 Oct;45(9):1043–9.

[46] Gläscher J, Adolphs R, Damasio H, Bechara A, Rudrauf D, Calamia M, et al. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proc Natl Acad Sci. 2012 Sep 4;109(36):14681–6.

[47] Sliwinska MW, Vitello S, Devlin JT. Transcranial Magnetic Stimulation for Investigating Causal Brain-behavioral Relationships and their Time Course. J Vis Exp. 2014 Jul 18;(89):51735.

[48] Goldstein WM, Hogarth RM. Judgment and decision research: Some historical context. 1997;

[49] Savage LJ. The foundations of statistics. Courier

Corporation; 1972.

[50] Lichtenstein S, Slovic P. The construction of preference. Cambridge University Press; 2006.

[51] Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. Nature. 1999 Jul;400(6741):233–8.

[52] Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. Science. 2003;299(5614):1898–902.

[53] Wu G, Gonzalez R. Curvature of the probability weighting function. Manag Sci. 1996;42(12):1676–90.

[54] Hertwig R, Barron G, Weber EU, Erev I. Decisions from experience and the effect of rare events in risky choice. Psychol Sci. 2004;15(8):534–9.

[55] Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature. 2006;442(7106):1042–5.

[56] Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. Proc Natl Acad Sci. 2018;115(27):E6106–15.

[57] Ludvig EA, Spetch ML. Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? PLOS ONE. 2011 Jun 1;6(6):e20262.

[58] Thompson WR. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika. 1933;25(3/4):285–94.

[59] Slivkins A. Introduction to multi-armed bandits. ArXiv Prepr ArXiv190407272. 2019;

[60] Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT press; 2018.

[61] Frank MJ, Seeberger LC, O'reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science. 2004;306(5703):1940–3.

[62] Barron G, Erev I. Small feedback-based decisions and their limited correspondence to description-based decisions. J Behav Decis Mak. 2003;16(3):215–33.

[63] Weber EU, Shafir S, Blais A-R. Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. Psychol Rev. 2004;111(2):430.

[64] Hertwig R, Erev I. The description–experience gap in risky choice. Trends Cogn Sci. 2009;13(12):517–23.

[65] Wulff DU, Mergenthaler-Canseco M, Hertwig R. A meta-analytic review of two modes of learning and the description-experience gap. Psychol Bull. 2018;144(2):140.

[66] Madan CR, Ludvig EA, Spetch ML. Comparative inspiration: From puzzles with pigeons to novel discoveries with humans in risky choice. Behav Processes. 2019;160:10–9.

[67] Lefebvre G, Lebreton M, Meyniel F, Bourgeois-Gironde S, Palminteri S. Behavioural and neural characterization of optimistic reinforcement learning. Nat Hum Behav. 2017;1(4):1–9.

[68] Ludvig EA, Madan CR, McMillan N, Xu Y, Spetch ML. Living near the edge: How extreme outcomes and their neighbors drive risky choice. J Exp Psychol Gen. 2018;147(12):1905–18.

[69] Jessup RK, Bishara AJ, Busemeyer JR. Feedback produces divergence from prospect theory in descriptive choice. Psychol Sci. 2008;19(10):1015–22.

[70] Erev I, Ert E, Plonsky O, Cohen D, Cohen O. From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. Psychol Rev. 2017;124(4):369.

[71] Addessi E, Beran MJ, Bourgeois-Gironde S, Brosnan SF, Leca J-B. Are the roots of human economic systems shared with non-human primates? Neurosci Biobehav Rev. 2020 Feb 1;109:1–15.

[72] Yamada H, Tymula A, Louie K, Glimcher PW. Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. Proc Natl Acad Sci. 2013;110(39):15788–93.

[73] Nioche A, Bourgeois-Gironde S, Boraud T. An asymmetry of treatment between lotteries involving gains and losses in rhesus monkeys. Sci Rep. 2019;9(1):1–13.

[74] Stauffer WR, Lak A, Bossaerts P, Schultz W. Economic choices reveal probability distortion in macaque monkeys. J Neurosci. 2015;35(7):3146–54.

[75] Chen X, Stuphorn V. Inactivation of medial frontal cortex changes risk preference. Curr Biol. 2018;28(19):3114-3122. e[4]

[76] Hayden B, Heilbronner S, Platt M. Ambiguity aversion in rhesus macaques. Front Neurosci. 2010;4:166.

[77] So N-Y, Stuphorn V. Supplementary eye field encodes option and action value for saccades with variable reward. J Neurophysiol. 2010;104(5):2634–53.

[78] Raghuraman AP, Padoa-Schioppa C. Integration of multiple determinants in the neuronal computation of economic values. J Neurosci. 2014;34(35):11583–603.

[79] Farashahi S, Azab H, Hayden B, Soltani A. On the flexibility of basic risk attitudes in monkeys. J Neurosci. 2018;38(18):4383–98.

[80] Ferrari-Toniolo S, Bujold PM, Schultz W. Probability Distortion Depends on Choice Sequence in Rhesus Monkeys. J Neurosci. 2019 Apr 10;39(15):2915–[29]

[81] McCoy AN, Platt ML. Risk-sensitive neurons in macaque posterior cingulate cortex. Nat Neurosci. 2005;8(9):1220–7.

[82] Hayden BY, Platt ML. Temporal Discounting Predicts Risk Sensitivity in Rhesus Macaques. Curr Biol. 2007 Jan;17(1):49–53.

[83] Long AB, Kuhn CM, Platt ML. Serotonin shapes risky decision making in monkeys. Soc Cogn Affect Neurosci. 2009 Dec;4(4):346–56.

[84] Watson KK, Ghodasra JH, Platt ML. Serotonin transporter genotype modulates social reward and punishment in rhesus macaques. PloS One. 2009;4(1).

[85] Hayden BY, Heilbronner SR, Nair AC, Platt ML. Cognitive influences on risk-seeking by rhesus macaques. Judgm Decis Mak. 2008;3(5):389.

[86] O'Neill M, Schultz W. Coding of Reward Risk by Orbitofrontal Neurons Is Mostly Distinct from Coding of

Reward Value. Neuron. 2010 Nov;68(4):789–800.

[87] Heilbronner S, Hayden BY, Platt M. Decision salience signals in posterior cingulate cortex. Front Neurosci. 2011;5:55.

[88] Kim S, Bobeica I, Gamo NJ, Arnsten AF, Lee D. Effects of α-2A adrenergic receptor agonist on time and risk preference in primates. Psychopharmacology (Berl). 2012;219(2):363–75.

[89] Heilbronner SR, Hayden BY. Contextual Factors Explain Risk-Seeking Preferences in Rhesus Monkeys. Front Neurosci [Internet]. 2013 [cited 2020 May 14];7.

[90] Xu ER, Kralik JD. Risky business: rhesus monkeys exhibit persistent preferences for risky options. Front Psychol [Internet]. 2014 Apr 24 [cited 2020 May 14];5.

[91] Smith TR, Beran MJ, Young ME. Gambling in rhesus macaques (Macaca mulatta): The effect of cues signaling risky choice outcomes. Learn Behav. 2017;45(3):288–99.

[92] Chambon V, Théro H, Vidal M, Vandendriessche H, Haggard P, Palminteri S. Choosing and learning: outcome valence differentially affects learning from free versus forced choices [Internet]. Neuroscience; 2019 May [cited 2020 May 14].

[93] Palminteri S, Lefebvre G, Kilford EJ, Blakemore S-J. Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. PLoS Comput Biol. 2017;13(8):e1005684.

[94] Farashahi S, Donahue CH, Hayden BY, Lee D, Soltani A. Flexible combination of reward information across primates. Nat Hum Behav. 2019 Nov;3(11):1215–24.

[95] Heilbronner SR, Hayden BY. The description-experience gap in risky choice in nonhuman primates. Psychon Bull Rev. 2016 Apr;23(2):593–600.

[96] Niv Y. Reinforcement learning in the brain. J Math Psychol. 2009;53(3):139–54.

[97] Daw ND, Doya K. The computational neurobiology of learning and reward. Curr Opin Neurobiol. 2006;16(2):199–204.

[98] Croson R, Sundali J. The Gambler's Fallacy and the Hot Hand: Empirical Data from Casinos. J Risk Uncertain. 2005 May 1;30(3):195–209.

[99] Blanchard TC, Wilke A, Hayden BY. Hot-hand bias in rhesus monkeys. J Exp Psychol Anim Learn Cogn. 2014;40(3):280–6.

[100] Sescousse G, Caldú X, Segura B, Dreher J-C. Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci Biobehav Rev. 2013 May;37(4):681–96.

[101] Amiez C, Neveu R, Warrot D, Petrides M, Knoblauch K, Procyk E. The Location of Feedback-Related Activity in the Midcingulate Cortex Is Predicted by Local Morphology. J Neurosci. 2013 Jan 30;33(5):2217–28.

[102] Hayden BY, Platt ML. Gambling for Gatorade: risk-sensitive decision making for fluid rewards in humans. Anim Cogn. 2009 Jan;12(1):201–7.

[103] Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. Curr Opin Neurobiol. 2012 Dec;22(6):1027–38.

[104] Barberis NC. Thirty Years of Prospect Theory in Economics: A Review and Assessment. J Econ Perspect. 2013 Feb;27(1):173–96.

[105] Markowitz H. The Utility of Wealth. J Polit Econ. 1952 Apr;60(2):151–8.

[106] Wood W, Neal DT. A new look at habits and the habit-goal interface. Psychol Rev. 2007 Oct;114(4):843–63.

[107] Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997;275(5306):1593–9.

[108] FitzGerald TH, Seymour B, Bach DR, Dolan RJ. Differentiable neural substrates for learned and described value and risk. Curr Biol. 2010;20(20):1823–9.

[109] Wimmer GE, Poldrack RA. Reward learning and working memory: effects of massed versus spaced training and post-learning delay period. bioRxiv. 2020;

[110] Abdellaoui M, Bleichrodt H, Kammoun H. Do financial professionals behave according to prospect theory? An experimental study. Theory Decis. 2013 Mar 1;74(3):411–29.

[111] Prelec D. The Probability Weighting Function. Econometrica. 1998 May;66(3):497

[112] Loomes G, Sugden R. Regret theory: An alternative theory of rational choice under uncertainty. Econ J. 1982;92(368):805–24.

[113] Quiggin J. Generalized expected utility theory: The rank-dependent model. Springer Science & Business Media; 2012.

[114] Niv Y, Edlund JA, Dayan P, O'Doherty JP. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. J Neurosci. 2012;32(2):551–62.

[115] Tobler PN, Christopoulos GI, O'Doherty JP, Dolan RJ, Schultz W. Neuronal Distortions of Reward Probability without Choice. J Neurosci. 2008 Nov 5;28(45):11703–11.

Box 1: description- and experience-based behavioural models

In the present box we sketch the formalisms standardly employed to explain and quantify risk preferences in description-based and experience-based decisions. Description and experience paradigms radically differ in how they model decision under risk. In the description domain, risk preferences are the *direct* result of subjective deformations of probabilities and outcomes that are explicitly stated. On the other side, in the experience domain there is no separate representation of outcomes' probabilities and no explicit deformation of outcomes' values. Consequently, risk preferences are the *indirect* result of the learning process that links past outcome information to subsequent to choices. Eventually, these two approaches lead to different explanations of risk-attitudes.

Risk preferences in description-based paradigms are commonly explained by prospect theory (PT). The expected value of k iterations of the same gamble X (which is random variable) is computed as follows:

$$E(X) = \sum_{i=1}^{k} p_i x_i \tag{1}$$

where x_i is the value of an individual outcome and p_i the objective probability of the outcome. PT states that the utility of an outcome, that is the subjective value $u(x_i)$, is non-linear and modulated by different parameters: α and β, that are the power to which respectively a positive or negative outcome are elevated, and λ the loss aversion coefficient. Thus, the PT utility function is defined as follows:

$$u(x_i) = \begin{cases} x_i^{\alpha} & if\ x_i \geq 0 \\ -\lambda(-x_i)^{\beta} & if\ x_i < 0 \end{cases} \tag{2}$$

an $\alpha \leq 1$ corresponds to risk-aversion in the gain domain (the intuition dates back to Bernoulli), $\alpha > 1$ corresponds to risk-seeking behaviors. In the loss domain, the same relation is true concerning the values of β. Value of $\lambda > 1$ corresponds to loss aversion; its typical empirical value is around 2 [10,110]. A decision-maker with $\alpha < 1$, $\beta > 1$ and $\lambda > 1$ will present different risk preference in the gain (risk aversion) and the loss (risk seeking) domain (**see Figure 3B**).
In addition, PT postulates a subjective deformation of probabilities. There are multiple ways to mathematically express the probability weighting function. One of the most common is the 'Prelec' function [111]:

$$\pi(p_i) = e^{-\delta(-\log(p_i))^{\gamma}} \tag{3}$$

with δ controlling the elevation, and γ the curvature. When both parameters are set to 1, the function tends to linearity. The more $\gamma > 1$, the more the function adopts an S-shape. A classical result is the overweighting of low probabilities compared to high probabilities, where the direction of the curve follows an inverse S-shape (see **Figure 3A**), with $\gamma < 1$. Note that other probability weighting function has been proposed [53]. Finally, the subjective expected utility is given by:

$$SEU(X) = \sum_{i=1}^{k} \pi(p_i) u(x_i) \tag{4}$$

By the variation of these parameters, PT allows to account for inter-individual differences in risk preferences. Of note, concurrent theories such as regret theory [112] or rank-dependent utility models [113], which use very different representational structures and parameterizations, are also used to model decision under risk.

Experience-based paradigms can be seen as reinforcement learning problems operationalized as k-armed bandit tasks [60]. Consider an environment composed by a state vector S, with $s \in S$. In each states s, there are available actions denoted $a \in A$. Each state-action pair has an underlying reward probability distribution, such that $P[R|s, a]$, is the probability of obtaining the reward R, knowing the state-action couple (s, a). An agent must then follow a policy in order to maximize a state-action value function $Q(s, a)$ (i.e. to maximize the average expected reward). A common learning policy is to compute subsequently to a choice the prediction error δ, that will be used to incrementally update the value associated to a specific state-action pair (s, a):

$$\delta = R - Q(s, a) \tag{5}$$

$$Q(s, a) \leftarrow Q(s, a) + \alpha \delta \tag{6}$$

with α the learning rate, that determines to what extent newly acquired information overrides previous one.

In this paradigm, inter-individual variability in behaviors can be accounted by differences in individual parameters such as the aforementioned learning rate α. However, this model with only one parameter is too simple to accommodate different risk preferences.

A way to refine this model to account for different risk preferences, is to allow for two different learning rates, α^+ and α^-:

$$Q(s,a) \leftarrow Q(s,a) + \begin{cases} \alpha^+\delta & if\ \delta > 0 \\ \alpha^-\delta & if\ \delta < 0 \end{cases} \tag{7}$$

If $\alpha^+ = \alpha^-$, the two learning rates model is equivalent to a one learning rate model. We define the tendency to preferentially update $Q(s,a)$ from positive prediction errors rather than negative prediction errors as *positivity bias* (or *loss neglect*) ($\alpha^+ > \alpha^-$). Conversely, we define the opposite situation ($\alpha^+ < \alpha^-$) as *negativity bias* (or *loss enhancement*).

The learning rate asymmetry has direct consequence for risk preferences in the setting where a subject has to learn the value of a safe (say a fixed value of 0) and a risky (say 50% chance of winning/losing one euro) option. A subject displaying an optimistic bias would neglect the past losses and will therefore be risk-seeker (see **Figure 3B**). Conversely, the pessimistic bias implies risk-aversion. Both pessimistic and optimistic biases have been reported in the literature, with the latter bias being more frequently reported [67,92,93,114].

While it is tempting to see the optimistic bias as the experience-based antithesis of loss aversion, their formalism and psychological interpretations are quite different and they are therefore not mutually exclusive. Indeed, loss aversion concerns the valuation of *prospective* losses, while the optimistic bias concerns the *retrospective* assessment of past losses.

It is important to note that, in humans, although the average values of the behavioural biases are reported as described above (for instance: inverse S-shape in description-based paradigms and loss neglect in experience-based paradigms; **see Figure3A**), there results are further tempered by high degree of inter-individual variability in the bias parameters. At the individual level, some subjects may in fact display opposite biases in both experimental settings [67,115]. If inter-individual variability is equally high in other primates, the fact that monkey studies use very small sample size (see **Figure 2**) can explain the comparably less consistent picture observed (see **Table 1**).