Specific Sensitivity to Rare and Extreme Events: Quasi-Complete Black Swan Avoidance vs Partial Jackpot Seeking in Rat Decision-Making*

Mickaël Degoulet† Louis-Matis Willem‡ Christelle Baunez§
Stéphane Luchini¶ Patrick A. Pintus‖

March 6, 2024

*This paper supersedes a previous version circulated under the title "Decision-Making in Rats is Sensitive to Rare and Extreme Events: the Black Swan Avoidance". The authors would like to thank, without implicating, Raouf Boucekkine, Andrea Brovelli, Giorgio Coricelli, John Duffy, Thibault Gajdos, Brian Hill, Tobias Kalenscher, Fabio Maccheroni, Stefano Palminteri, Drazen Prelec, Guillaume Rocheteau and Wolfram Schultz for very useful discussions and encouragement at early and more recent stages of this research, which has had a long gestation lag, as well as many conference and seminar participants for helpful feedback, especially at the UC Irvine Econ department, the MUSEES conference in Lyon, the Comparative Psychology Research Colloquium in Düsseldorf, LNC2 in Paris, IAST in Toulouse.

†Institute of Neurosciences Timone, UMR7289 CNRS & Aix-Marseille Université. Email: mickael.degoulet@univ-amu.fr.

‡Ecole Centrale Marseille. Email: louis.willem@centrale-marseille.fr.

§Institute of Neurosciences Timone, UMR7289 CNRS & Aix-Marseille Université. Email: christelle.baunez@univ-amu.fr.

¶Aix-Marseille University, CNRS, AMSE. Email: stephane.luchini@univ-amu.fr.

‖CNRS-Humanities and Social Sciences and Aix-Marseille University, CNRS, AMSE. Email: patrick.pintus@univ-amu.fr.

Abstract: Most studies assessing animal decision-making under risk rely on probabilities that are typically larger than 10%. To study Decision-Making in uncertain conditions, we explore a novel experimental and modelling approach that aims at measuring the extent to which rats are sensitive - and how they respond - to outcomes that are both rare (probabilities smaller than 1%) and extreme in their consequences (deviations larger than 10 times the standard error). In a four-armed bandit task, stochastic gains (sugar pellets) and losses (time-out punishments) are such that extremely large - but rare - outcomes materialize or not depending on the chosen options. All rats feature both limited diversification, mixing two options out of four, and sensitivity to rare and extreme outcomes despite their infrequent occurrence, by combining options with avoidance of extreme losses (Black Swans) and exposure to extreme gains (Jackpots). Notably, this sensitivity turns out to be one-sided for the main phenotype in our sample: it features a quasi-complete avoidance of Black Swans, so as to escape extreme losses almost completely, which contrasts with an exposure to Jackpots that is partial only. The flip side of observed choices is that they entail smaller gains and larger losses in the frequent domain compared to alternatives. We have introduced sensitivity to Black Swans and Jackpots in a new class of augmented Reinforcement Learning models and we have estimated their parameters using observed choices and outcomes for each rat. Adding such specific sensitivity results in a good fit of the selected model - and simulated behaviors that are close - to behavioral observations, whereas a standard Q-Learning model without sensitivity is rejected for almost all rats. This model reproducing the main phenotype suggests that frequent outcomes are treated separately from rare and extreme ones through different weights in Decision-Making.

Introduction

Although exploration and exploitation are key drivers of observed behavior when humans and other animals interact repeatedly with their environment (see e.g. [26] for an overview), not much is arguably known about how they are affected by outcomes that are both very rare and very consequential, that is, *Rare and Extreme Events* (REE thereafter). While risk is a central dimension in the related and enormous literature that contributes to social sciences and neurosciences, as it should, most studies on Decision-Making rely on lotteries for which the frequencies of outcomes are typically larger than 10% (e.g. [9], [10], [37], [49]; see also [19] for exceptions when choices are, however, not repeated). This is admittedly a

major limitation in view of the fact that many living species experience events that are not only much rarer but that also entail potentially very large -, negative as well as positive - impacts. The subsequent evolutionary consequences for - and possibly the extinction of - large groups or species constitute an important example (see e.g. [18]). On the brighter side, human evolution has been punctuated by the design and spread of artefacts, a few of which have allowed big leaps forward on the way to material comfort (see e.g. [27]).

To the extent that both their consequences and their frequencies are highly uncertain, because historical data either lacks or is scarce, such REE are challenging for Decision-Making. Too little is known about whether animal species, including humans, are sensitive to - and how they cope with - such extremely infrequent and consequential events when making choices. In this study, we focus on rats, which have been formerly used to assess Decision-Making processes in various behavioral tasks (e.g. [6], [1], [45], [47], [23], [30], [43], [20], [54]; see also [22] on the relevance of animal models for human Decision-Making), but mostly because they allow further neurobiological manipulations in future studies. We have designed a novel experiment in which rats interact repeatedly with their laboratory environment through a 4-armed bandit task. In this task, they face REE that occur with a frequency smaller than 1% and, equally importantly, that are associated with extreme outcomes in terms of *both* gains and losses - deviations that are larger than 10 times the standard error. Following the design of the Iowa Gambling Task for human participants (see [4]), related tasks for rodents typically assume that the less frequent events have probability values that are about or larger than 10% (see e.g. [1], [34] and [53]).

In addition to dealing with low-probability and high-consequence events, the key feature of our experimental design is that it allows to test whether, in such a context, sensitivity to REE favors choosing frequently exposures to uncertainty that limit the consequences of extremely "negative" risks - or losses - while leaving open the possibilities to benefit from "positive" and large risks - or gains (called "Anti-fragile" exposures in [48]). Symmetrically, therefore, an option that rules out extreme gains but may trigger very large losses (labeled Black Swans) is labeled Fragile in our design, and it would be chosen frequently or exclusively when sensitivity to REE is rather low or nonexistent. More precisely, complete exposure to Jackpots and complete *avoidance* of Black Swans both define the Anti-fragile option, while the reverse configuration is representative of the Fragile one (i.e. complete avoidance of Jackpots and complete exposure to Black Swans, rather oddly). We also consider Robust and Vulnerable options: the former, which may seem to correspond to

resilience, rules out exposure to any REE while the latter exposes to both REE, positive and negative. As such, our approach is of interest in an evolutionary perspective, which would ask how often the Anti-fragile option is indeed selected by animals and humans, and whether there are natural foundations for Anti-fragile choices that possibly relate to fitness and are embodied in the brain.

Importantly, our experimental paradigm distinguishes three domains as follows: "normal events" (NE), which occur with approximately 90% probability, "rare events" (RE) happening with about 10% probability, and REE which occur with a probability smaller than 1%. Each option is defined by a relationship between the magnitudes of rewards and the frequencies at which they occur. In, say, the gain domain, while one of the two options dominates the other for NE, this dominance weakens when RE are added to to picture and is eventually reversed if REE occur. To put it simply, the option that is dominated for NE and RE but turns out to dominate when REE occur is such that the less frequent the reward, the more consequential it is in terms of larger values: we therefore label this option convex. It follows that convex options expose to Jackpots and, by the same token, avoids Blacks Swans while, symmetrically, concave options avoids Jackpots and expose to Black Swans. Therefore, the Anti-fragile option is convex because it combines accelerating and possibly large gains (labeled Jackpots thereafter), together with decelerating and limited losses, when probabilities decrease - that is, when the number of trials needed to observe such gain occurrences increase. Here, to repeat, convexity does not refer to the shape of utility or other value functions, but refers to the fact that the *magnitude* of gains/losses, is accelerating/decelerating against decreasing probability (i.e. increasing number of trials). In sum, our experimental design is based on the assumptions that: 1) with perfect information about the probability distributions of all exposures, value-maximizing participants whose preferences are represented by any non-decreasing value function, would choose concave gains in the domain restricted to NE, due to first-order *stochastic dominance*. In other words, in the NE domain, it is more interesting to choose the concave options; 2) they would continue to do so in the RE domain for any non-decreasing and concave value function, due to second-order Stochastic Dominance; 3) however, in the full domain with REE, value-maximizing subjects endowed with any non-decreasing and concave value function are predicted to choose convex gains, because of a reversal in second-order Stochastic Dominance. In other words, exposure to Jackpots and avoidance of Black Swans is by design costly and disadvantageous when REE do not materialize.

Because of our specific interest in REE on both positive (rewards) and negative (losses) values, our framing integrates gains (sugar pellets) and losses (time-out punishment preventing the rats to earn pellets). One attractive feature of our design is that it delivers two direct and fairly natural measures that help interpreting behavioral data. First, *Total Sensitivity to REE* measures the extent to which rats take into account REE by combining convex options with avoidance of extreme lossees (Black Swans) and exposure to extreme gains (Jackpots). Formally, Total Sensitivity adds up the fractions of convex choices in terms of both gains and losses. Second, *One-sided Sensitivity to REE* captures choices that preferably, and asymmetrically, favor either the seeking of large gains (Jackpot Seeking) or the avoidance of large losses (Black Swan Avoidance). That is, One-sided Sensitivity is defined as the difference between the fraction of convex choices in gains and that in losses. While the Anti-fragile option has by design maximal Total Sensitivity and the Fragile option has zero Total Sensitivity, they both exhibit zero One-Sided Sensitivity. In contrast, the other alternatives, i.e. the Robust and Vulnerable options, have both middle Total Sensitivity but differ in that they respectively exhibit REE avoidance (negative One-Sided Sensitivity) and REE seeking (positive One-Sided Sensitivity). Of course, rats have the possibility to combine all four options over the course of the experiments.

Using our setup on a sample of 20 rats, with about 6000 stimuli per rat over 41 sessions, we document two sets of behavioral results. First, all rats diversify their choices but in a limited fashion across the possible set of options, primarily 2 out of the 4 available. Most rats (19 out of 20) exhibit moderate to high levels of Total Sensitivity. This means that most rats choose to diversify their choices across options that often combine some exposure to extreme gains and some avoidance of extreme losses. More strikingly, most rats (13 out of 20) tend to exhibit *quasi-complete* Black Swan avoidance mixed with *partial* Jackpot seeking: they have high Total Sensitivity to REE and favor avoiding almost certainly Black Swans even though this entails less frequent exposure to Jackpots (negative One-Sided Sensitivity to REE), by mixing mostly the Anti-fragile and Robust options. We interpret such a behavior as evidence of the *differential treatment* of rare and extreme gains and losses. Consistent with such an interpretation, we also find that Total Sensitivity and Black Swan avoidance averaged over all sessions are reinforced after rare and extreme losses are effectively experienced in the trials.

We next examine the behavioral observations though the lens of Reinforcement Learning models, which are theoretical benchmarks in the exploration-exploitation literature (see e.g.

[46]). More specifically, the modelling contribution of this study is to introduce sensitivity to Black Swans and Jackpots in a new class of augmented Q-Learning models and we estimate their parameters using observed choices and outcomes for each rat. Interestingly, the model selected through information criteria has a distinctive feature: it separates NE/RE from REE in the decision to pick one option rather than any of the other. In other words, the selected model says that most rats behave as if they attach different weights to normal/rare outcomes and to rare and extreme outcomes, especially losses, in Decision-Making.

Results

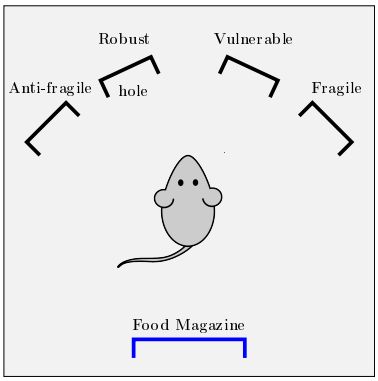
Measures of sensitivities to REE in a four-armed bandit task

Figure 1 summarizes the main features of our experimental setting. In panel (a) is depicted a schematic representation of the Skinner box, in which all rats performed all 41 experimental sessions . Each session consisted of about 120 trials. Each rat had the possibility at all times to poke into 4 holes, each corresponding to a different sequence of gains and losses, drawn at random for each rat from a fixed distribution across sessions. Gains were sugar pellets while losses were time-out punishments (time period during which nosepoke remained inactive, thus entailing an opportunity cost in terms of pellets not consumed) measured in seconds (see the Material and Methods sections on Experimental Model and Subject Details, Experimental Method Details). In the following we use "options" for possible choices and "choices" for observed choices. The key feature of the four possible options associated to each available hole was that they were combinations of convex (that is, accelerating) and/or concave (that is, decelerating) gains and losses with decreasing frequencies of occurrence, in panel (b) of Figure 1 (see the Material and Methods section on Modelling and Statistical Analysis for the Jensen gap as a quantitative measure of convexity). While, as shown in the first and last columns, values for losses and gains differed for convex and concave options, the associated probabilities were assumed to be identical. They were set, as seen in the middle column, to $\{p_1 = 0.25 - \varepsilon/3, p_2 = 0.2 - \varepsilon/3, p_3 = 0.05 - \varepsilon/3, \varepsilon\}$ for both concave and convex gains. Parameter ε is the (very small) ex-ante probability of the largest outcome that we label the *Jackpot* for the 80-pellet gain, and the *Black Swan* for the 240-second loss. Both outcomes are denoted Rare and Extreme Events (REE in short).[1]

[1] As a convenient shortcut, therefore, the Black Swan is thought of as a rare and extreme loss, to differentiate it from a rare and extreme gain.

Figure 1: Experimental design - panel (a), schematic representation of the rat behavioral operant chamber equipped with four holes, each associated with a specific exposure to gains and losses (Anti-fragile, Robust, Vulnerable and Fragile); panel (b), values and frequencies for convex (green)/concave (red) gains/losses, with the largest probabilities highlighted in blue for Normal Events (20 to 25%, NE domain), Rare Events in yellow (5%, RE domain) and Rare and Extreme Events in pink (ε-probability smaller than 1%, REE domain); panel (c), representation and label of the four possible options, with increasing gains/losses (up/down) against decreasing probabilities (right/left): "Robust" has convex losses (avoidance of negative REE called Black Swan) and concave gains (avoidance of positive REE called Jackpot), "Anti-fragile" has convex losses and gains (Black Swan avoidance and Jackpot seeking), "Fragile" has concave losses and gains (Black Swan seeking and Jackpot avoidance) and "Vulnerable" has concave losses (Black Swan seeking) and convex gains (Jackpot seeking); panel (d), representation of two measures of Sensitivity to REE, i.e. "Total Sensitivity" (y-axis) and "One-sided Sensitivity" (x-axis) and examples, where green dot represents sensitivities associated with an hypothetical near-Anti-fragile profile (close to full Jackpot seeking and complete Black Swan avoidance), red dot represents sensitivities associated with an hypothetical near-Fragile profile (close to full Black Swan seeking and Jackpot avoidance), and blue dot represents sensitivities associated with an hypothetical choice mixing mostly Anti-fragile and Robust (near complete Black Swan avoidance and partial Jackpot seeking) - see the Material and Methods section on Modelling and Statistical Analysis for details

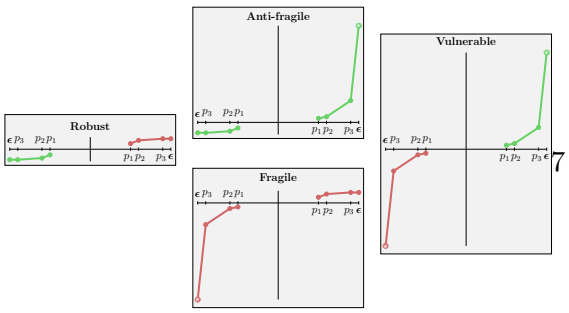
(a)



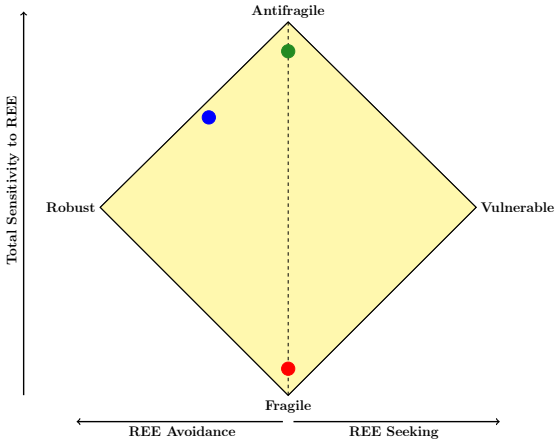
(b)

Losses (sec)			Gains (pellets)	
convex	concave	Probability	convex	concave
6	3	≈0.25	1	2
12	9	≈0.20	3	4
15	36	≈0.05	12	5
15	240	ϵ	80	5

(c)



(d)



The next largest values (that is, 12 pellets and 15 seconds, vs 5 pellets and 36 seconds) are labeled a Rare Event (RE), since their total probability is around 10% and the associated values are moderate compared to REE. Note that for concave gains and convex losses the RE and REE values coincide. This is to contrast with the Jackpot and Black Swan that occur only if convex gains and concave losses are chosen, respectively. With total probability about 90%, therefore, the remaining set of lowest values in the first two rows (highlighted in blue in panel (b) of Figure 1) for concave and convex options are labeled Normal Events (NE), and these are the typical events most considered in the literature. Combining gains and losses that are either convex or concave then delivers 4 options.

The 4 options that correspond to the 4 holes in the Skinner box are pictured in panel (c) of Figure 1, with the following meaning. On the horizontal x-axis are reported (in decreasing order moving away from the origin) the ex-ante probabilities that are unknown to the subject, which only observes the outcome that is measured on the vertical y-axis.[2] For convenience, gains appear in the upper-right quadrant while losses appear in the lower-left one. For example, the option named *Anti-fragile* (after [48]) at the top of panel (c) is composed of convex gains and convex losses: the right part depicts the points in the (x, y) axis are $\{p_1, 1\}, \{p_2, 3\}, \{p_3, 12\}, \{\varepsilon, 80\}$, corresponding to the fourth column in panel (b); the left part corresponds to the points $\{\varepsilon, -15\}, \{p_3, -15\}, \{p_2, -12\}, \{p_3, -6\}$, where "negative" values are interpreted as losses in seconds. Looking at the green curves that summarize the Anti-fragile option, then, one visualizes convexity directly from the properties that gains accelerate moving right from the origin, while losses decelerate moving left from the origin, with decreasing probability. Alternatively, decreasing probability can be thought of as increasing numbers of trials that are needed to experience the associated gain or loss. As a consequence, the Anti-fragile option potentially exposes to the Jackpot but avoids exposure to the Black Swan.[3]

Symmetrically, the *Fragile* option in the lower part of panel (c) in Figure 1 has both gains and losses that are concave (in red). Gains are $\{p_1, 2\}, \{p_2, 4\}, \{p_3, 5\}, \{\varepsilon, 5\}$, which

[2] In contrast with [10], we do not give rats any cue about the events' frequencies during the training sessions, which are set by us as experimenters. This makes our setting ecologically closer to experience-based designs.

[3] While in panel (c) convex (concave) gains and losses are identified by color, in green (red), a geometric description might be useful as well. For example the property that a monotone, upward or downward, convex (concave) curve lies below (above) its chords joining any pair of points can be used to identify green (red) curves.

means that gains plateau at 5 pellets, compared with 80 for the Jackpot attached to convex gains in the Anti-fragile option. In the loss domain, the Fragile option delivers $\{\varepsilon, -240\}, \{p_3, -36\}, \{p_2, -9\}, \{p_3, -3\}$: it exposes to the Black Swan with an awfully long delay of 240 seconds. For the Fragile option, then, gains decelerate while losses accelerate, with decreasing probability, so that it potentially exposes to the Black Swan but avoids exposure to the Jackpot. Finally, the ways the *Robust* and *Vulnerable* options combine convex/concave gains and losses may be viewed as symmetric: the Robust choice protects from the Black Swan but at the same time misses the opportunity to get the jackpot, while the Vulnerable option potentially delivers both REE since both gains and losses accelerate with decreasing probability (or increasing number of trials).

In sum, the four options in panel (c) of Figure 1 are therefore combinations of the convex/concave gains/losses in panel (b): the Anti-fragile option at the top is convex both in gains and in losses, while the Robust option (left) is convex only in the loss domain. On the other hand, the Vulnerable exposure (right) has convex gains only, while the Fragile option (bottom) is concave both in gains and in losses. This implies that the Jackpot of 80 pellets may materialize only when either the Anti-fragile or the Vulnerable exposure are picked, while the Black Swan of 240 seconds may be experienced only by rats choosing either the Fragile or the Vulnerable exposures. In the Material and Methods section on Modelling and Statistical Analysis, it is shown that convex gains/losses dominate concave gains/losses in the sense of Stochastic Dominance, provided that the probability of REE ε is larger than about 0.3%. In our design, the values for ε we assumed to meet this condition but still are below 1%, which admittedly ensures that REE are indeed rare. In addition, we provide in the Material and Methods section on Modelling and Statistical Analysis, a characterization of the convexity properties of all options, in terms of Jensen gaps that we relate to statistical moments in closed-form.

Panel (d) in Figure 1 is the central graphic tool for presenting the main behavioral data from our experiments. Because, as emphasized so far, convexity of available options is a key dimension in our design about decision-making under REE, our first task is to track how often rats have chosen convex rather than concave options in gains/losses, over the course of the 41 sessions that each of our 20 rats has run. Second, we also aim at assessing whether rats have possibly picked convex options more or less often in the loss domain than in the gain domain, which would depict an asymmetric sensitivity to REE. Both dimensions easily combine in panel (d). On the vertical y-axis we report the *sum* of the frequencies of

choosing both convex gains (that is, Anti-fragile or Vulnerable) and convex losses (that is, Anti-fragile or Robust), while on the horizontal x-axis is depicted the *difference*, defined as the frequency of picking convex options in the gain domain minus the frequency of choosing convex options in the loss domain. The resulting rotated square in panel (d) of Figure 1 is then a convenient way to report each rat's choices over one or several sessions, along both dimensions and can be used to identify different types of behavior. More details about how to formally construct such a rotated square are given in Material and Methods, subsection Modelling and Statistical Analysis.

To see why panel (d) is useful, let us first focus on its 4 edges, which represent extreme cases in the sense that they correspond to specialized choices: if an hypothetical rat's behavior is summarized by a point located at exactly the Anti-fragile edge, it means that this rat has exclusively chosen the Anti-fragile option at each nose poke during all 41 sessions. As a consequence, the vertical y coordinate reaches its maximum value (2 if frequencies are added up) while the horizontal x coordinate is zero since both convex gains and convex losses are - symmetrically - chosen all the time. Following the same logic, the Fragile edge has coordinate $\{0, 0\}$ since it involves choosing concave gains and losses all the time, i.e. never choosing convex options. The Robust option, in contrast, is depicted by $\{-1, 1\}$ since it implies picking exclusively convex losses but concave gains, in an asymmetrical fashion. The reverse asymmetry characterizes the Vulnerable edge at $\{1, 1\}$, with convex gains and concave losses chosen at all times. Now when a rat's behavior happens to be represented by a point that is *not* located at any of the 4 edges, such a location indicates that that rat diversifies (that is, mixes) across options. For example, the green point in panel (d) is close to the Anti-fragile edge and located on the vertical line linking it to the Fragile edge, which might happen if a rat would pick the former with, say, frequency about 95% of the time and the latter about 5% of the time. The red point would be attained by reversing those proportions. Of course many other combinations of the 4 options is possible. For example, the blue point depicts an hypothetical rat that would diversify across the Anti-fragile and Robust options almost equally, at the expense of the two remaining options that have negligible frequencies. Th next section will show that diversifying primarily across two options is indeed the main pattern featured in our behavioral data.

A feature of our experimental design and its implications need to be emphasized in more details at this point. First, the set of NE has total probability of about $2(p_1+p_2) \approx 0.9$ since ε is a small number. RE are rare but not extreme is the sense that their total probability

$2p_3 \approx 0.1$ is moderately small while their consequences are moderately large. REE, in contrast, are more extreme outcomes, since they imply waiting 240 seconds (the Black Swan) and gaining 80 pellets (the Jackpot), with a much smaller likelihood. In practical terms, the frequency of each REE is ex-ante around 1% for all rats. Despite their small probability, REE have some importance in the following sense. As shown in Material and Methods, subsection Modelling and Statistical Analysis, the values for both frequencies and outcomes ensure that concave options dominate, in the sense of first-order stochastic dominance, convex options over NE, as well as in the sense of second-order stochastic dominance over the union of NE and RE. In other words, if the extreme gains and losses do not materialize, concave gains (losses) dominate convex gains (losses). However, this does not necessarily hold when REE are added: provided that $\varepsilon \geq 0.302\%$ (which we assume hold ex-ante), convex gains (losses) dominate concave gains (losses) in the sense of second-order stochastic dominance. This implies, in particular, that over the full domain that is the union of NE, RE and REE, expected gains for convex options are larger than expected gains for concave options. Likewise, expected time-out punishments for convex options are smaller than expected gains for concave options over the full range including NE, RE and REE.

The reason that we impose such a dominance reversal is as follows. In theory, a rat choosing concave options at all times can be thought of as having absolutely no sensitivity to REE: it acts as if those events never occur and always go for first or second-order stochastic dominance (over NE and RE) as Decision-Making criteria. And those criteria unequivocally point to the Fragile option as the best choice. This gives our first measure along the y-axis of panel (d), Total Sensitivity to Rare and Extreme Events, that again simply sums up the proportion of convex exposures that are chosen for each rat over the 40 sessions. A rat picking the Fragile option at all times has zero Total Sensitivity, since it behaves as if the Black Swan or the Jackpots never occur. At the other extreme, a rat picking exclusively the Anti-fragile option signals maximal Total Sensitivity, since it always stays exposed to the Jackpot while always avoiding the Black Swan. Moderate Total Sensitivity arises, then, if a rat diversifies its choices, for instance primarily across either Anti-fragile and Fragile, or Robust and Vulnerable. Note that such pairs obviously achieve identical outcomes in terms of pellets and time-out punishments when chosen with equal frequencies.

Because gains and losses are deliberately integrated in our design, we also need to distin-

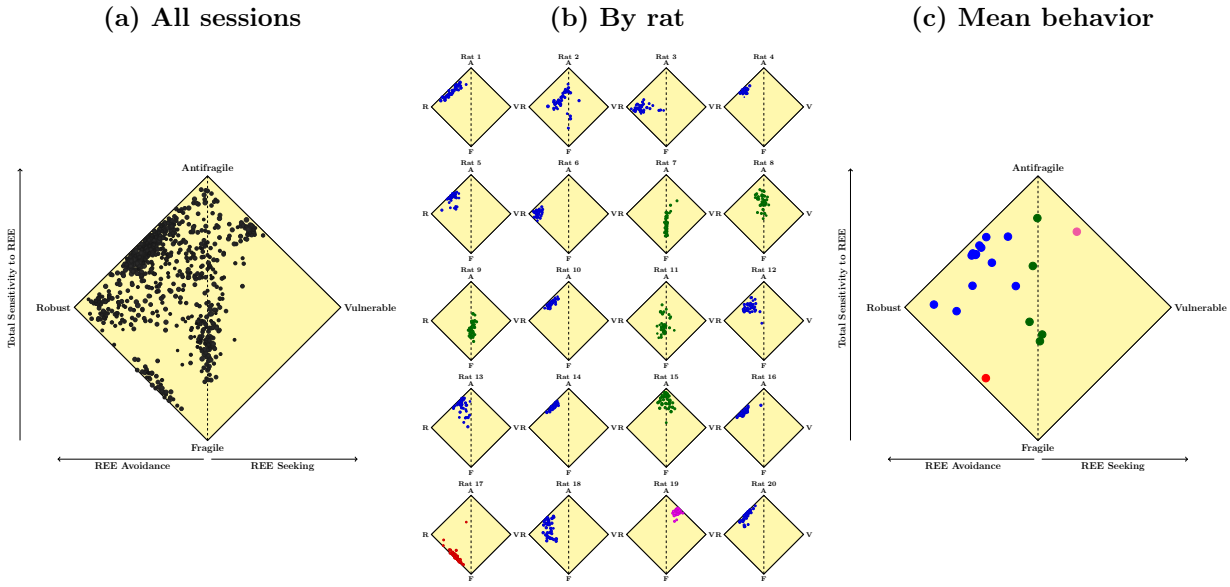
guish whether rats tend to choose convex exposures symmetrically over the gain and loss domains. This is the purpose of the measure along the x-axis in panel (d). Suppose that we focus for now on the upper half of the square, which implies that Total Sensitivity to REE is larger than one. We then say that a particular rat exhibits Black Swan Avoidance when it picks convex exposures in the loss domain more often than in the gain domain. Symmetrically, Jackpot Seeking is the label we use to depict a situation when convex choices are more often made in the gain domain. Note that these statements, rather qualitative at this stage, will be quantified below. Given its Total Sensitivity larger than one, a rat's behavior exhibits Black Swan Avoidance (Jackpot Seeking) whenever it is represented by a point that is located on the upper right (left) quarter of the square in panel (d). Symmetrically, for Total Sensitivity smaller than unity, one concludes to Jackpot Avoidance (Black Swan Seeking) whenever the rat is represented by a point that is located on the lower right (left) quarter of the square In other words, the numbers read on the y-axis measure *One-Sided Sensitivity to Rare and Extreme Events*: when negative (positive) it indicates REE Avoidance (REE Seeking). Fully tooled-up with both measures, we report in the following sections our main results.

Main phenotype features quasi-complete Black Swan Avoidance but partial Jackpot Seeking

In this section, we present the behavioral data coming from the 41 sessions that we have been running for each of the 20 rats - see the Material and Methods section on Experimental Method Details for more details. We organize the data in the following way, so as to underline two features of the choices made by the sample of rats. In each panel of Figure 2 is depicted the (rotated) square that we have already presented in panel (d) of Figure 1, the edges of which represent the four exposures if chosen exclusively at all times. In the square, each point represents a rat. A quick glance at Figure 2 already shows that virtually no such point is located exactly at any of the 4 edges. Quite to the contrary, since it turns out that *all rats diversify their choices across a set of options*, which is reminiscent of observed behaviors such as, for example, bet-hedging in animals and financial portfolio strategies used by humans (see [5], [36], [39], [44]).

How rats diversify their choices can be represented in three steps, focusing first on all choices made by all 20 rats over the 41 sessions, depicted by black points in panel (a) of Figure 2. More precisely, each black dot represents a rat's session, with a total of about 800

Figure 2: Total Sensitivity to REE (y-axis) against One-sided Sensitivity to REE (x-axis; i.e. REE seeking towards the right and avoiding towards the left) - see Material and Methods section on Modelling and Statistical Analysis - panel (a) black-colored points depict 41 sessions run by 20 rats all together; panel (b) in each square representing a rat, colored points depict 41 sessions for each of the 20 rats; the color code indicates the 2 main categories of profiles observed (blue: the most sensitive to REE and black swan avoiders; green: sensitive to REE but neither biased towards seeking nor avoidance (close to midline on the graph); 2 outliers were identified and color-coded in red and pink; panel (c) color-coded points depict averages over the 41 sessions for each of the 20 rats



such points. Most points lie on the right half of the square, that is, on the REE Avoidance side. In addition, within the right half most points are located in the upper quarter, that is, above the Robust-Vulnerable vertex, suggesting moderate to high Total Sensitivity as well as Black Swan Avoidance. In the Material and Methods section on Experimental Method Details, we discuss the contribution of the training sessions to such patterns. These two patterns are better seen if we focus on each individual rat.

We now look separately at individual behavior for each of the 20 rats across the 41 sessions, in panel (b) of Figure 2. Each rat is represented by a square and, within each square, a point represents one of the 41 sessions for that particular rat. For convenience, the vertices of all squares are labelled by letters: "A" obviously stands for the Anti-fragile

exposure, while "R", "V" and "F" stand respectively for the Robust, Vulnerable and Fragile options. All rats exhibit moderate to high Total Sensitivity, with the exception of rat 17 depicted in red in panel (b) of Figure 2. Among this vastly dominant group, one can identify two sub-groups, color-coded in blue and green, and a second outlier. Starting with the latter, rat 19 (in purple) is the sole rat that exhibits both the highest Total Sensitivity and REE (Jackpot) Seeking, by mixing mostly Anti-fragile (quite often) and Vulnerable (less often) options.

Except for rats 17 and 19, the remaining 18 rats can be gathered in two sub-groups. The largest one gathers the rats colored in blue: rats 1 to 6, 10, 12 to 14, 16, 18 and 20 diversify their choices primarily across the Robust and Anti-Fragile options, with various frequencies. This means that *the dominant phenotype, observed over 13 rats depicted in blue, exhibits both moderate to high Total Sensitivity together with Black Swan Avoidance*: rats within this group try more often to avoid the Black Swan than to be exposed to the Jackpot. A smaller sub-group composed of the rats depicted in green, however, diversify their choices rather across the Fragile and Anti-Fragile options: *those 5 rats depicted in green exhibit also Total Sensitivity but feature neither Black Swan Avoidance nor Jackpot Seeking*: those rats have zero One-Sided Sensitivity or, in other terms, make choices that feature symmetry towards in terms of neither avoiding nor seeking REE.

Third, the two groups of blue and green rats are conveniently represented in panel (c) of Figure 2. Each point is associated to a particular rat and it represents the *average* of its choices over the 4A sessions, as measured through our two metrics. As for Total Sensitivity, one notices that all rats but 4 have an average sensitivity larger than half its maximal value: 3 rats having a Total Sensitivity slightly below 1 and a single rat with a value around 0.5. The latter (rat 17 in red) is the closest to a Fragile behavior, with a significant deviation from it, though. Overall, therefore, most rats (that is, 19 out of 20) are rather sensitive to the presence of REE. In addition, panel (c) confirms the striking feature that most rats exhibit Black Swan Avoidance (i.e. negative One-Sided Sensitivity). Rats 17 and 19 (in red and purple respectively) are outliers for opposite reasons. The latter is the unique rat with low Total Sensitivity while the former is the rat with the largest Total Sensitivity and with pronounced Jackpot Seeking behavior. Among the remaining group composed of 18 rats, all with moderate to high Total Sensitivity, the average behaviors in panel (c) allow one to identify the two sub-groups that we have previously stressed: the 5 rats depicted in green have zero One-Sided Sensitivity (neither Black Swan Avoidance nor Jackpot Seeking) while

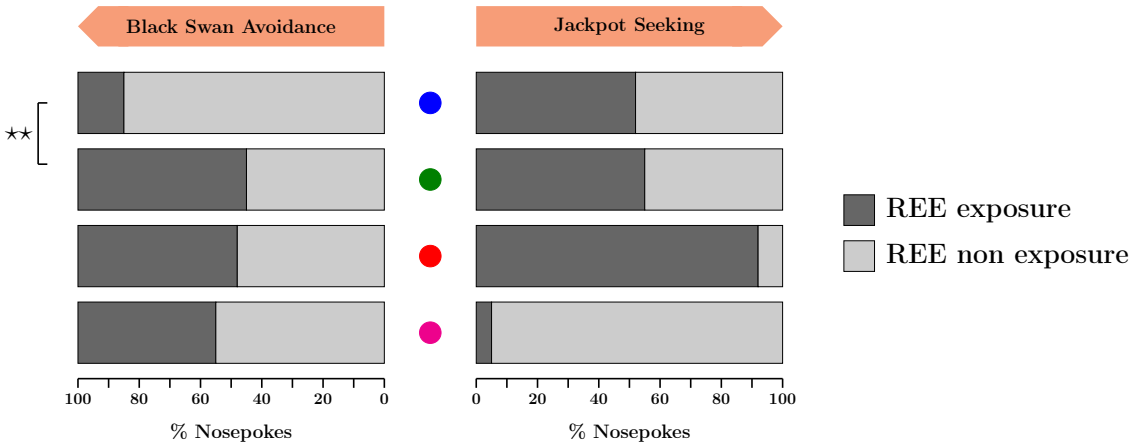
the 13 blue rats exhibit pronounced Black Swan Avoidance. Exact values can be found in the Supplementary Material section on Behavioral measures for each rat.

An interesting characteristic appears from direct inspection of panel (c) in Figure 2: the sample of 16 rats with the largest Total Sensitivity features *a positive relationship between Total and One-Sided Sensitivities*. In other words, among those rats that have both moderate to high Total Sensitivity and Black Swan Avoidance, rats that are more sensitive overall to convexity tend to also exhibit lower Black Swan Avoidance (that is, less negative One-Sided Sensitivity). It is as if large Total Sensitivity comes with a mixture of exposures that tend to favor convexity in the gain domain relative to the loss domain. In a sense, rat 19 in purple can be seen as a limiting example of that behavior: it combines the Anti-fragile and Vulnerable options so as to approach near-maximal Total Sensitivity and Jackpot Seeking as opposed to Black Swan Avoidance.

This property is confirmed if we compute the correlation between average Total Sensitivity and average One-Sided Sensitivity. For example, for the whole sample of 20 rats, this correlation is close to (and not statistically different from) zero, while it has a value of about 0.57 over the 41 sessions within the group of 16 rats with Total Sensitivity larger than one. In other words, if we drop the 4 rats with the smallest Total Sensitivity from the sample (those rats represented by points below the Robust-Vulnerable vertex in panel (c) in Figure 2), we are left with 16 rats for which larger Total Sensitivity is associated with lower One-Sided Sensitivity/less Black Swan Avoidance, that is, more symmetry and even Jackpot Seeking for rat 19. Moreover, if we divide the 41 sessions of the 16-rat group over time, the correlation between Total and One-Sided Sensitivity increases from about 0.17 on average in the first 10 sessions, to 0.56 in the next 10 sessions, to 0.71 in the next 10 sessions, and finally to 0.64 in the final set of 10 sessions. Therefore, the pattern we alluded to appears early in the sessions and is consistent over time.

In Figure 3, we additionally report behavioral measures of Black Swan Avoidance and of Jackpot Seeking, defined as the shares of nosepokes that respectively lead to exposure to Jackpots and non-exposure to - i.e. avoidance of - Black Swans (how those behavioral measures relate to Total and One-Sided Sensitivities is described in details in the Material and Methods section on Modelling and Statistical Analysis). What the median values of Black Swan Avoidance and of Jackpot Seeking that appear in Figure 3 reveal is that the blue and green phenotypes differ sharply in one respect: half of the rats depicted in blue achieve nearly complete Black Swan Avoidance since more than about 90% of their

Figure 3: Median behavioral measures of Black Swan Avoidance and Jackpot Seeking by phenotype, defined as median shares of nosepokes exposed to REE (in percentage) for blue and green phenotypes; phenotypes are indicated in column with a corresponding colored symbol at the center - left panel depicts the percentage of nosepokes exposed to Black Swans in black and not exposed to Black Swans in light gray - Black Swan Avoidance increases with the share of non exposure to Black Swans; right panel depicts the percentage of nosepokes exposed to Jackpots in black and not exposed to Jackpots in light gray - Jackpot Seeking increases with the share of exposure to Jackpots; blue and green phenotype comparisons are carried out by Wilcoxon Rank Sum tests ($\star$ means $p < 0.05$ and $\star\star$ means $p < 0.01$)

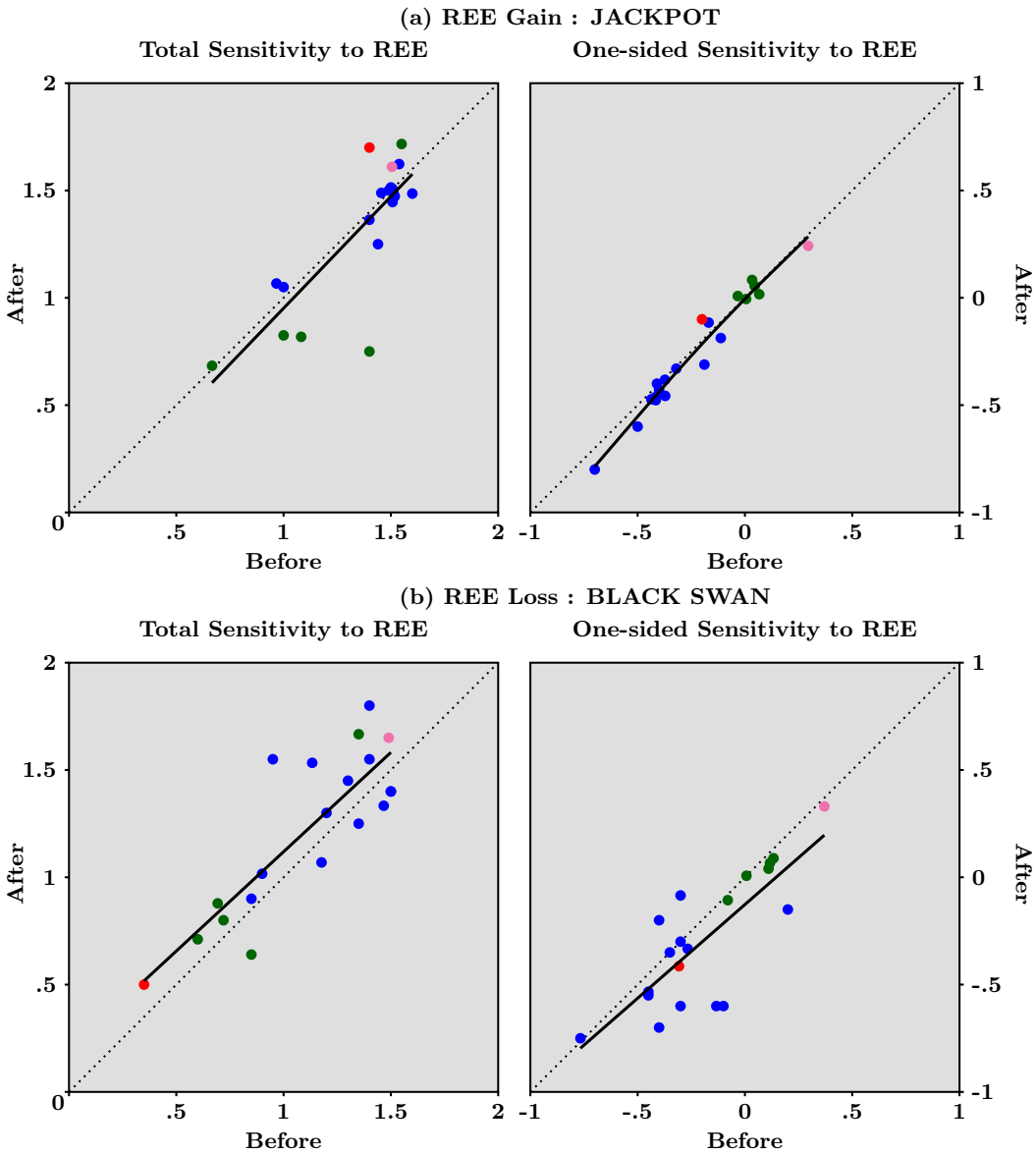


nosepokes lead to avoid being exposed to Black Swans, compared to more than about 40% for half of the green rats as well as for the red rat and pink rat. In contrast, blue and green rats do not differ in terms of median Jackpot Seeking. This is confirmed by paired Wilcoxon tests: Black Swan Avoidance and Jackpot Seeking are significantly different for blue rats ($p = .0002$) but not for green rats ($p = .6250$). The two rats classified as outliers (red and pink) differ from both blue and green rats in terms of Jackpot Seeking. The red rat seeks quasi-complete exposure to Jackpots whereas the pink rat seeks quasi-complete avoidance of Jackpots. Exact values can be found in the Supplementary Material section on Behavioral measure for each rat.

Black Swans do reinforce sensitivities to REE while Jackpots do not

Having documented that most rats (19 out of 20) exhibit medium to high Total Sensitivity and medium to high Black Swan Avoidance, we now address the issue of how rats *respond*

Figure 4: Responses of Total Sensitivity to REE and of One-sided Sensitivity to REE, averaged over the 41 final sessions, following Jackpots in panel (a) and Black Swans in panel (b), for each of the 20 rats. Each dot indicates one rat and dots are color coded to indicate the profile of the animal, as in Figure 2. Black dotted lines materialize no change in sensitivities after exposure to the REE compared to before. Black solid lines represent spline estimates - see Material and Methods section on Modelling and Statistical Analysis



to the actual occurrence of REE within a session. More precisely, we aim at comparing the average behavior of each rat before and after the occurrence of Black Swans and Jackpots. In Figure 4, we report how REE affect Total and One-Sided Sensitivities, within a window formed by the average over 10 nose pokes before and the average over 10 nose pokes after. For both panels (a) and (b), each point represents a particular rat and the colors group all rats as in panels (b)-(c) of Figure 2. In panel (a) of Figure 4 we report the impact of Jackpots while in panel (b) is depicted the effect of Black Swans, for Total Sensitivity on the left and One-Sided on the right. Dotted 45-degree lines represent the no-effect benchmark.

Visual inspection of both panels suggests that while Jackpots tend not to affect both sensitivities much, Black Swans do and tend to increase Total Sensitivity and to decrease One-sided Sensitivity: most dots do tend to be located above (below) the 45-degree line in the lower left (right) panel. In other words, rare and extreme losses/Black Swans tend to reinforce Total Sensitivity, contrary to extreme gains/Jackpots. This reinforcement effect due to experiencing negative REE is consistent with our main result that most rats are sensitive and chose to avoid extreme losses, that is, exhibit medium to high Total Sensitivity and Black Swan Avoidance. In addition, panel (b) also reveals that rare and extreme losses tend to make Black Swan Avoidance stronger: most dots tend to be located below the 45-degree line, while in the lower right panel. Again, this is consistent with the observation that extreme losses are somewhat given more weights than extreme gains by most rats, which implies that Black Swan Avoidance is reinforced when a extreme loss is experienced. This is confirmed by statistical tests, including one using bootstrap, as shown in the Material and Methods section on Modelling and Statistical Analysis (see in particular the discussion of Table 3). In sum, a within-session analysis of the effects of experiencing a REE shows that the occurrence of Black Swans tends, in the short-run, to significantly reinforce both Total Sensitivity and Black Swan Avoidance, while the occurrence of Jackpots tend to leave Total and One-Sided Sensitivity unaltered.

Q-Learning models require specific REE decision weights to mimic rats' behavior

We have introduced sensitivity to Black Swans and Jackpots in a new class of augmented Reinforcement Learning models and we have estimated their parameters using observed choices and outcomes for each rat. The selected model ended up with a distinctive feature: it separates normal from rare and extreme outcomes through different weights in the Decision-

Making process. Adding such specific sensitivity results in a good fit of the selected model - and simulated behaviors that are close - to behavioral observations, whereas a standard Q-learning model without sensitivity to REE is rejected for almost all rats.

To each of the four available options, indexed below by "o" (from 1 to 4, referring to Antifragile, Fragile, Robust, Vulnerable) in equations (1), we attach at each moment in trial/time t a gain sub-value $Q^g(o)$ and a loss sub-value $Q^l(o)$ that are updated as follows:

$$\begin{aligned} \underbrace{Q^g_{t+1}(o) - Q^g_t(o)}_{\text{update of gain sub-value}} &= \underbrace{\alpha_g}_{\text{gain learning rate}} \times \underbrace{[r^g_t(o) - Q^g_t(o)]}_{\text{gain prediction error}} \\ \underbrace{Q^l_{t+1}(o) - Q^l_t(o)}_{\text{update of loss sub-value}} &= \underbrace{\alpha_l}_{\text{loss learning rate}} \times \underbrace{[r^l_t(o) - Q^l_t(o)]}_{\text{loss prediction error}} \end{aligned} \tag{1}$$

where the convention that updated values are indexed by $t+1$ means right after observing the gain r^g_t or loss r^l_t corresponding to the choice made at t. Parameters α_g and α_l are usually labeled learning rates, indicating the speed at which reward prediction error gets updated into the latest values. Limiting cases occur when setting the learning rate either to zero - its smallest possible value - which means no subvalue updating, or to one - its largest possible value - which means that the subvalue tracks the obtained reward itself. While equations (1) applies for the updating of the chosen option, a similar updating rule applies for the three other options that are not chosen at that moment, by assuming that this happens as if the reward was zero, with different parameters. That is, for options that are not chosen, the updating rules for subvalues are simply $Q^g_{t+1}(o) = (1-\alpha^f_g)Q^g_t(o)$ and $Q^l_{t+1}(o) = (1-\alpha^f_l)Q^l_t(o)$, where the forgetting rates α^f_g and α^f_l are assumed to be bounded between zero and one.

The novel feature of our augmented Q-Learning model, besides integrating gains and losses, is that specific weights are attached to options that may produce REE. More specifically, we model the value of each option o as:

$$\underbrace{V_t(o)}_{\text{value of option}} = \underbrace{\lambda_g Q^g_t(o) + \lambda_l Q^l_t(o)}_{\text{decision weight on sub-values}} + \underbrace{\gamma_g 1_{JP}(o) + \gamma_l 1_{BS}(o)}_{\text{decision weight on REE}} \tag{2}$$

In equation (2), the value of each option V_o is the sum of two terms. The first is the decision weight attached that sums up the gain and loss subvalues, each weighted by parameters λ_g and λ_l that may reflect a differential effect. Importantly, in the class of nested models that we estimated, REE may or may not be incorporated in the gain and loss subvalues, that

Table 1: Selected augmented Q-Learning models by phenotype

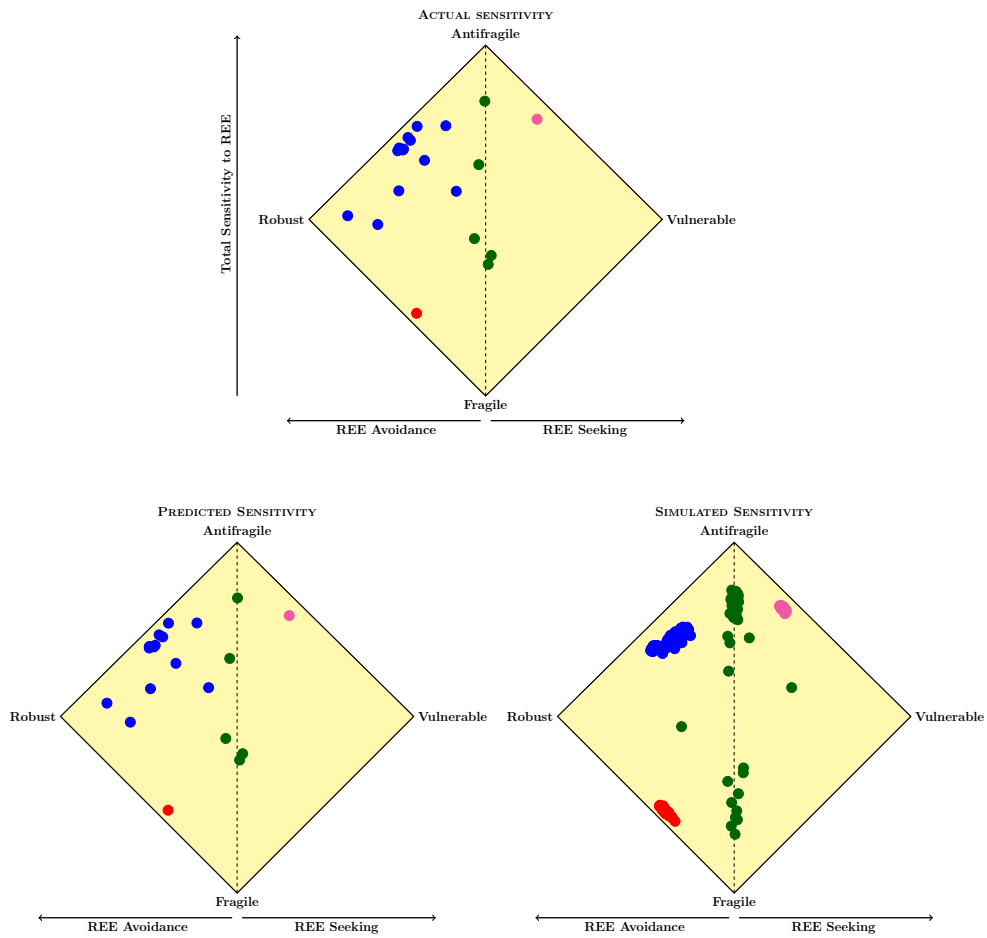
Model selected	Rats ♯ by phenotype
No integration of REE (distinct forgetting parameters)	• $\{3; 12\}$ • $\{11\}$
Only decision weights on REE	• $\{2; 5\}$ • $\{17\}$ • $\{19\}$
Only decision weights on REE (distinct forgetting parameters)	• $\{1; 6; 10; 13; 14; 18\}$ • $\{7; 8; 9; 15\}$
Decision weights on REE and REE in Q sub-values	• $\{4; 16; 20\}$

should be thought of as averages - though not arithmetic but of an exponential moving type in view of equations (1) - see page 32 in [46]. The second and key term in equation (2) specifically captures the decision weight attached to REE, and it can itself be decomposed into the decision weights on Jackpots and on Black Swans. More precisely, the indicator function $1_{JP}(o)$ equals one if the option exposes to Jackpots (namely when the chosen option is either A or V) and zero otherwise. Similarly, $1_{BS}(o)$ equals one if the option exposes to Black Swans (namely options F and V) and zero otherwise. Therefore, γ_g and γ_l reflect the (possibly different) weights attached to Jackpots and Black Swans in the decision to choose one particular option rather than any other available.

Finally, the associated so-called action probabilities are given, for each option, by the Softmax function (a.k.a. the multinomial Logit model) $\pi_t(o) = e^{V_t(o)}/\sum_{i=1}^{4} e^{V_t(i)}$, each a number between zero and one. In the estimation procedure detailed below, the parameters of interest are γ_g and γ_l in the sense that the benchmark model sets them to zero. Our candidate models are then compared against such a benchmark in two ways: one sets of models incorporates REE into sub-values while the second does not, and γ_g and γ_l are estimated for both. We hope it has become clear by now that the interest of the latter set of models is of interest because it indicates, if selected, that REE are integrated into the Decision-Making process *in a separate way and not through averages*, that is, not through the sub-values. In sum, the class of augmented Q-Learning models with gains and losses has at most 8 parameters: two learning rates, two forgetting rates, two decision weights for sub-values and two decision weights for REE.

We have used information criteria to determine the selected model for each rat over all its choices, as summarized in Table 1 by phenotype. Selected models indicate that 17 rats out of 20 integrate REE, at least using specific decision weights, in their Decision Making process, while only the remaining three do not. Most animals (8 blue rats, 4 green rats, as well as pink and red rats) integrate REE only though their specific decision weights (with or

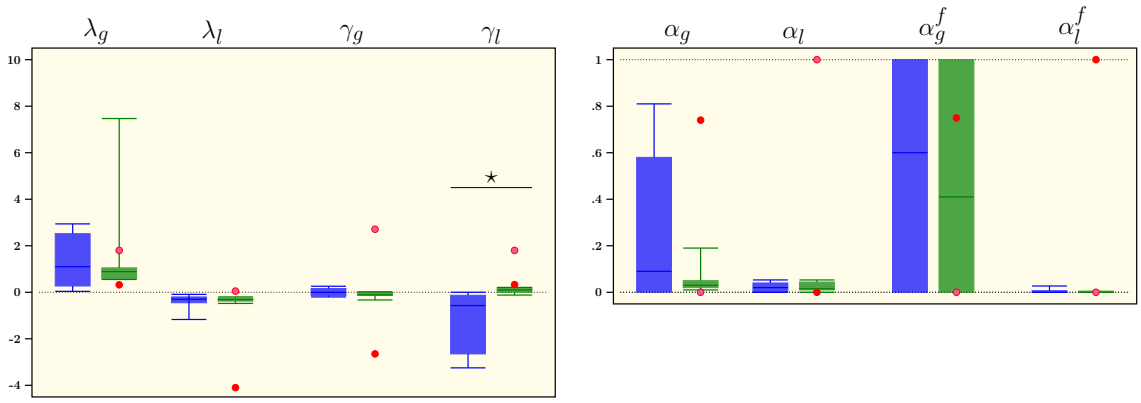
Figure 5: Total Sensitivity to REE (y-axis) against One-sided Sensitivity to REE (x-axis) - averages for each rat over the 41 sessions; top panel replicates the right panel in Figure 2: observed sensitivities; bottom left panel: predicted sensitivities derived from selected models; bottom right panel: simulated sensitivities computed from runs of selected models over artificial 41 sessions



without specific forgetting parameters). Three blue rats integrate REE both with specific decision weights and in Q-sub-values. In contrast, the learning and forgetting parameters do not help to discriminate models.

Figure 5 reports the predicted (bottom left) and simulated sensitivities (bottom right), which can be compared to actual sensitivities (top, which replicates the right panel in Figure 2). The predicted sensitivities, generated using the estimated parameters of the selected model given the actual rewards obtained during the 41 sessions, are accurate descriptions of the observed sensitivities computed from the actual choices. Simulated sensitivities

Figure 6: Parameters of the selected Q-Learning model, estimated for each rat over the 41 sessions - left panel: decision weights for averages of gains and losses and for REE; right panel: learning and forgetting rates for gains and losses; while individual parameter for pink and red rats are represented as dots, parameters for blue and green rats are represented by box plots with 10th and 90th percentiles, interquartile ranges and median; between phenotype comparisons carried out by Wilcoxon Rank Sum tests ($\star$ means $p < 0.05$) - see Material and Methods subsection Augmented Q-Learning Model Estimation and Simulation for details



are obtained by running the selected model over artificially generated 41 sessions, using its median parameter values for blue and green rats (and individual parameter values for the pink and red rats). These simulations show first that the selected Q-Learning model generate behavioral data that are stable for the blue rats as well as for the red and pink ones. Second, simulated green rats tend to separate into two stable groups, still along the vertical line.

Finally, estimated parameter values appear in Figure 6, color-coded in the same way as in previous figures. The details about nested models and how the selected one is obtained are presented in the section of the Materials and Methods on Augmented Q-Learning Model Estimation and Simulation. The main result in Figure 6 relates to a major difference between the behaviors of blue and green rats, as seen through the lens of our augmented Q-Learning model: blue rats have a significant specific sensitivity to Black Swans that is translated into a very negative decision weights attached to options that expose to them ($p = 0.0102$). In other words, blue rats have a negative γ_l, which means that they tend to avoid the options with Blacks Swans (F and V) most of the time. Along that dimension, they differ significantly from green rats, which mostly mix options F and A.

Discussion

Quasi-complete Black Swan Avoidance vs partial Jackpot Seeking as evidence that extreme gains and extreme losses are treated differently

To sum up the behavioral results: we have shown, first, that all rats in our sample exhibit some Total Sensitivity - in fact, medium to large Total Sensitivity for most rats - to REE. And, second, that the main phenotype features nearly complete Black Swan Avoidance coupled with partial Jackpot Seeking. We interpret these results as intimately related to a fundamental asymmetry between negative and positive REE. In our design, rats have to endure larger frequent delays to avoid Black Swans and to forgo larger frequent gains for a Jackpot to occur. However, rats can avoid with certainty the Black Swans when doing so, making convex menus over losses relatively more attractive, but cannot guarantee that a Jackpot will occur, making convex options over gains relatively less attractive. The actual occurrence of the latter will depend on the particular draw obtained in each of the repeated sessions. As an illustrative example, suppose one uses a Poisson approximation of the probability that the REE materializes given that 150 choices are made within a session, which falls within the range of the number of nose-pokes that rats have performed. If one further supposes that the ex-ante probability of REE is 1%, then the probability to observe exactly one REE in that session is about 78%, which leaves some room for the REE to *not* occur. Our design thus captures an asymmetry that arguably distinguishes positive from negative REE in general: while it may be possible to completely avoid the latter, the former is never expected to occur for sure even if the appropriate choices are repeated. In fact, the rarer the Jackpot, the larger the number of trials needed to experience at least one, requiring marked perseverance in Jackpot Seeking.

In the specific context of our experimental task, this asymmetry implies that although complete Black Swan Avoidance means avoiding Black Swans with certainty, complete Jackpot Seeking does not ensure that Jackpots will necessarily occur. In sharp contrast, the values corresponding to the frequent domain - i.e. the union of NE and RE - will be observed often given their ex-ante probabilities larger than 10%. In other words, while inference about how often REE - and Jackpots for that matter - occur is not reliable, accurate inference about how the frequent gains materialize seems much more within reach given the many choices made in the conditioning and final sessions.

This feature of the experimental design has important consequences for the choices that rats make. To the extent that they explore different options, rats are subjected to a flood of stimuli that point at concave options as the optimal choice: the Fragile option, in particular, generate larger gains but smaller losses in the frequent domain and in fact it stochastically dominates all the others if REE do not materialize. To put it differently, for each session, the Fragile option is dominated by the Anti-fragile one if both the Black Swan and the Jackpot occur. However, because complete Black Swan Avoidance practically insures against the occurrence of Black Swans, it seems like an attractive combination even if this implies violating Stochastic Dominance in the frequent domain: Stochastic Dominance over the full domain is attained with certainty. In contrast, complete Jackpot Seeking is less attractive since it does not ensure that a Jackpot will occur: the cost of getting smaller gains most of the time means that partial Jackpot Seeking makes sense as a good compromise since it implies that Stochastic Dominance in the frequent domain is less often violated, in a context where Stochastic Dominance over the full domain cannot be attained with certainty.

In sum, bearing the cost of getting smaller gains frequently is acceptable only if the compensating benefit in the form of a Jackpot materializes when many choices are made. Similarly, bearing the cost of getting larger losses frequently is acceptable only if the compensating benefit in the form of an avoided Black Swan is ensured. While the latter condition is certainly met if Black Swan Avoidance is complete, the former condition is not even if Jackpot Seeking is complete. Results suggest that rats depicted in blue have learned this property, since they combine mostly the Robust and Anti-fragile options to obtain quasi-complete Black Swan Avoidance together with partial Jackpot Seeking. In contrast, the rats depicted in green seem not to have learned this property, as they mix Anti-fragile and Fragile options, thus remaining exposed to Black Swans.

In theory, since frequent outcomes occur often when choices are repeated many times, Stochastic Dominance in the frequent domain should be detected and learned rather quickly by animals. However, the experimental literature shows that animals need a large number of repeated sessions to learn Stochastic Dominance (e.g. [55], [3]). We think our results suggest a potential explanation : animals may need a significant amount of sampling to infer that REE are excluded. In addition, one may wonder at this stage whether our results can be expressed in terms of loss - and possibly ambiguity - aversion. Although most other studies have not focused on REE, to the best of our knowledge, neuroeconomic tasks using positive and negative outcomes have generally shown that loss aversion occurs

(see e.g. [17], [13], [57]). This bias towards negative/losses in the valuation of outcomes has been previously shown and is in line with Prospect Theory ([21]), although models following this approach fit better rat's behavior in tasks only involving positive rewards and in which loss is not a punishment, but rather a lack of positive reward ([10],[3]). It could also include some emotional/anxiety-related component ([56]), possibly involving the basolateral amygdala as shown in rats ([50]). Ambiguity aversion in animals, on the other hand, seems to have been less explored (see [40] for non-human primates). Our results tentatively suggest that rats might conform to ambiguity *reduction* regarding REE, when combining Black Swan Avoidance that is quasi-complete and Jackpot Seeking that is only partial: while the Jackpot's sampling frequency of occurrence is in each session bounded below by zero and above by ε at best, the Black Swan's sampling frequency approaches zero when Black Swan Avoidance is nearly complete.

The quasi-complete Black Swan Avoidance together with partial Jackpot seeking, that we report for the main phenotype represented by rats depicted in blue, provides further evidence that *extreme* gains and losses receive differential treatment in the Decision-Making process. However, we show that the classical Q-Learning models which account for a potentially different treatment of gains and losses do not fit our behavioural data. Interestingly enough, observed behaviors are in contrast neatly captured in the selected Q-Learning models that are augmented with specific sensitivities to REE, and precisely through a marked avoidance to Black Swans which sets the main phenotype represented by rats depicted in blue apart from rats depicted in green.

The augmented Q-Learning models suggest a specific cerebral pathway for REE

Both the good fit and the accurate simulations of the augmented Q-Learning model for the main behavioral phenotype suggest a novel neurobiological conjecture. Viewed from the lens of the model, 11 rats (out of the 13 that form the blue group, as detailed in the Material and Methods section) exclude REE for the Q sub-values while expressing a specific sensitivity to Black Swans through a negative parameter γ_l. This means that when evaluating available options, "blue" rats do not incorporate Black Swans into the average of losses but rather attribute them a specific decision weight, with the effect that options exposing to Black Swans become less attractive. For humans, extreme stimuli have been shown to be more psychologically salient in perception and memory than moderate

stimuli in [25]. Our experimental and modelling results suggest that something similar could happen for rats. This insight from modelling opens up the possibility that rare and extreme outcomes, especially rare and extreme losses, could be encoded differently in the brain, compared to frequent and smaller rewards that are classically considered to be encoded by the DA neurons ([9], [42], [57], [51]), the striatum ([7], [11]), and cortical areas such as cingulate and orbitofrontal cortex ([41], [14])

Further studies are necessary to identify how REE could be encoded and where. If we consider that taking REE into account would require to integrating long-term outcomes and not only short term considerations as studied by [52], we might hypothesize that the regions of interest could be the anterior insula and amygdala. The latter being well known for its involvement in emotional processes, it is wise to consider it could be involved in REE-related events, which are highly salient and emotionally charged, and affect Decision-Making with a differential treatment of negative outcomes ([7], [50]). From a statistical viewpoint, it makes sense not to add up "outliers" into the first moment - the mean - of the observed distribution, because doing so could result in an unstable sample average. In a related way, good Decision-Making might depend on giving special treatment to REE over frequent rewards - as well as differential treatment of Black Swans vs Jackpots - and if such practice is evolutionary advantageous in uncertain environments (see e.g. [8]), this opens up the possibility that this would be reflected in the brain, possibly in different species. To the best of our knowledge, such an hypothesis that builts upon experimental and modelling results such as the ones we report here has not been put to the test in neurosciences, as it should in view of the arguments developed above for humans and other animals. This aspect could in principle be addressed within a computational validity framework, as advocated in [38].

Relevance for humans

Whether humans would find Black Swan Avoidance and Jackpot Seeking attractive and how they would combine them in a related experimental setup is a question that comes naturally to mind in view of the results reported in this paper. Our novel experimental design can in fact be adapted to human participants, so as to test in a similar fashion whether they adopt behavior that tend to avoid harmful - and seek beneficial - REE. This makes the present results obtained with rats also relevant for pressing questions that link the past and future of humanity. Usually gathered under the concept of the Anthropocene,

there is mounting evidence that climate and other environmental changes mainly driven by human activities originate extreme events (see e.g. [12]), raising the possibility of rare ones that could lead to partial or full extinction for many species. Even though sizable uncertainty remains about the magnitudes involved, the associated empirical evidence puts the human species in a specific situation that raises difficult but pressing questions. Why has the human species not succeeded in the past in avoiding courses of action that could potentially lead to such harmful REE, and to what extent will humans be able to cope with them in the near future, should they happen? In other words, are humans engaged in a path that diverges from Black Swan Avoidance instead of approaching it?

While intuitive, the notion that some adaptive and survival value derives from animal behaviors that help avoid destructive REE, or at the very least help cope with them should they happen, still lacks compelling evidence (see, however, [8]). If correct, such a conjecture would raise a paradox, given the specific contribution of humans in creating environments that are more prone to harmful REE: do humans possess a special trait that somehow prevented them to insure against such REE, despite the evolutionary advantage possibly attached to avoiding them? Put differently, to the extent that destructive REE are less exogenous to humans, due to their unique set of technological capabilities, than they are to other species, could there exist an epigenetic rule that is specifically human and that is conducive to behaviors that are unable to rule out extremely harmful events? If that is the case, could such an epigenetic rule manifest itself in such a way that the capacity of humans to cope with such REE, if they have become unavoidable, is hampered? For example, is it the case that detecting somewhat predictable cues for REE, such as convexity in our setting, is out of reach (see [2])?

In this study, we report results obtained from behavioral experiments conducted on rats that do not go, fortunately, in the direction of a positive response to the above set of questions. More precisely, the *heterogeneity* in individual behavior that we observe in rats, when confronted with alternative options in an uncertain context, suggests that some animals may individually lack the ability to insure themselves against a rare but consequential loss, even if they have the possibility to do so. In that sense, the human species, if afflicted by the same incapacity, is not alone in the animal kingdom. Of course, the bright side of the coin is that *most* rats consistently chose options that avoid almost completely exposure to a harmful REE and even provide partial exposure to beneficial REE in the form of a large gain. As a matter of fact, this group is the most numerous in our

sample of rats that have run the experiments. A speculative hypothesis then comes to mind: that some human individuals, if not all, may also settle on behaviors that avoid harmful - and possibly seek beneficial - REE. More generally, if the sensitivity and reaction to REE differ among individuals in some species, this raises the important question of how such heterogeneity affects - and interacts with - population dynamics and social institutions, most importantly collective Decision-Making.

Material and Methods

Experimental Model and Subject Details

Animal Subjects

Adult Lister Hooded males (n = 20, 200 g at arrival, Charles River) were housed in groups of two in Plexiglas cages and maintained on an inverted 12 h light/dark cycle (light onset at 7 pm) with water available *ad libitum*, in a temperature - and humidity - controlled environment. Food was slightly restricted (80% of daily intake). Animal care and use conformed to the French regulation (Decree 2010-118) and were approved by local ethic committee and the French Ministry of Agriculture under #03129.01.

Experimental Method Details

Apparatus

All behavioral experiments took place during the animals' dark phase in standard five-hole operant boxes (MedAssociates) located in ventilated sound-attenuating cubicles. One side of each box was equipped with a central house light, a tone generator and a food magazine, outfitted an infrared beam for detecting nose poke inputs. Sucrose pellets (20 mg; Bio-Concept Scientific) were delivered from an external food pellet dispenser. An array of five response holes was located on the opposite curved wall, each equipped with stimulus lights and infrared beams for detecting input (nose poke). The center hole was continuously closed throughout the experiments (Fig. 1a). Data were acquired on a PC running MedPC-IV.

Design of Behavioral Sequences

Four menus were elaborated by mixing convex and concave exposures for both the gains (sugar pellets) and the losses (time-out punishment) described on Figure 1 (panel (*b*)). For the gains, animals could obtain 1 , 3 (NE domain, blue), 12 (RE domain, yellow) or 80

(REE domain; pink) sugar pellets in the convex exposure or 2, 4, 5 or 5 pellets in the concave exposure. For the losses , convex exposure may impose 6, 12 , 15 or 15 sec of time-out punishment while it was 3, 9 (NE domain; blue), 36 (RE domain, yellow) or 240 sec (REE; pink) for the concave exposure.

The four behavioral options depicted in Figure 1, panel (c), are therefore combinations of the above convex and concave exposures: the "Anti-fragile" exposure at the top middle is convex in both gains and losses, while the "Robust" option (left) is only convex for the losses. On the other hand, the "Vulnerable" exposure (right) is convex only for the gains, while the "Fragile" option (bottom middle) is concave for both gains and losses. This implies that the extreme - but rare - gain of 80 pellets (i.e. Jackpot) may be delivered only when either the Anti-fragile or the Vulnerable options are picked, while the extreme and rare time-out punishment of 240 sec (i.e. Black Swan) may be only experienced if choosing either the Fragile or the Vulnerable options.

The first three events of each behavioral options belong to the frequent domain as their frequency of occurrence, respectively 0.5, 0.4 or 0.1, is significantly larger than zero. During behavioral testing, animals equivalently experienced the gain and loss domains. On the other hand, extreme outcomes, i.e. Jackpot or Black Swan, have a much smaller likelihood of occurring since they may appear only at particular point during the behavioral sequences (see below). This means that they could happen if a rat has chosen an exposure that is either convex in the gain domain or concave in the loss domain at a given time. This implies that the frequency of rare and extreme events is less than half of one percent for all rats. Despite their low probability, extreme events have some importance because of their value. Our calibration of both frequencies and outcomes ensures that concave exposures dominate convex exposures: if the extreme gain (Jackpot) does not materialize, the expected payoff in sugar pellets is larger for concave exposures than convex exposures (i.e. ex-post first order stochastic dominance, over the frequent domain). Similarly, the expected time-out punishment is lower for concave exposures compared to convex ones, if the Black Swan does not happen. However, ex-post first-order stochastic dominance is reversed in the presence of extreme events, in which case convex exposures become more interesting in terms of payoff than concave ones. The reason we imposed such a dominance reversal was as follows. If rats always choose concave exposures, then they show no sensitivity to rare and extreme events, since they act as if those events never occur and always go for first-order stochastic dominance. This gives us our first measure, Total Sensitivity to REE, that

simply sums up the proportion of convex exposures that are chosen for each rat over the 41 sessions. Because we deliberately integrate gains and losses in our design, we need to distinguish whether rats tend to choose convex exposures symmetrically over the gain and loss domains. We say that a particular rat exhibits Black Swan avoidance when it picks convex exposures in the loss domain more often than in the gain domain. Likewise, Jackpot Seeking occurs when convex choices are more frequent in the gain domain.

To avoid potential learning of the event occurrence during behavioral training and testing, ten different sequences of events, with respect of the first-order stochastic dominance as well as the balance between gain and loss domain exposures, were generated and randomly used for behavioral training and testing (Supplemental Figure S3). Furthermore, to increase the rarity and the unpredictable nature of extreme events, the sequences of events used during behavioral training and testing were declined into seven various sequences, in which Jackpot and Black Swan are either unavailable, solely or both available at a given time point of the sequence of events (Supplemental Table S1): when available, extreme events could be obtained at the 10th or the 60th activation within a given sequence, but could not occur at the same time. For example, in a behavioral sequence where the Jackpot should be available at the 10th position, any poke in the Anti-fragile and Vulnerable holes following nine responses made in one of these menus (regardless of the positive or negative outcomes) would trigger the delivery of the Jackpot. Of note, depending on the sequence used, animals could experience both extreme events in a single session.

Behavioral Training and Testing

Training was divided into five distinct phases before the final test: acquisition of the food collecting responses, acquisition of nose poking in the holes, training with four holes, attribution of menus to hole and training on the menus (Figure 7). Each session started with the illumination of the house light.

1. Acquisition of the food collecting responses: Animals were trained to collect sucrose pellets in the food magazine during three 30-min daily sessions (100 pellets max) under fixed ratio 1 schedule of reinforcement (FR1): one nose-poke into the food magazine triggered the delivery of one sucrose pellet. During this initial phase, nose-poke in any hole had no programmed consequences.

2. Acquisition of nose-poking in the holes: Here, animals had only access to one hole (the three others being occluded) during the 20-min sessions. One nose-poke in the avail-

able hole triggered the illumination of the hole-light and the delivery of one sucrose pellet in the food magazine. Perseverative nose-pokes (those performed before food collection) had no consequences. Following food collection in the food tray, animals were allowed to poke again in the opened hole. During each session, a maximum of 100 pellets were delivered. All animals were trained twice on each hole.

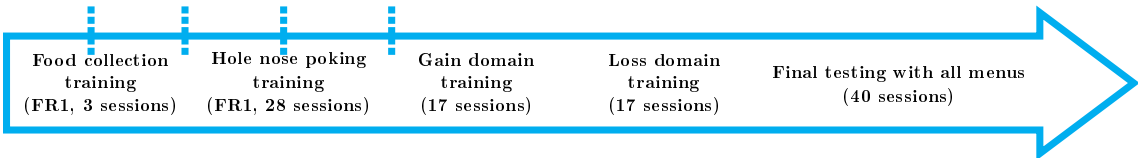
3. Training with four holes: Following the eight training sessions, animals were allowed to poke in the four different holes during twenty daily sessions of 20-min. Nose poke in any hole triggered the illumination of the associated light and the delivery of one sugar pellet in the food magazine. Both perseverative activations and pokes in other holes had no consequences. After food collection in the magazine, animals were allowed to nose poke again in any hole. The first ten training sessions were limited to 100 sugar pellets. During the following ten sessions, animals were able to collect up to 200 pellets per session.

4. Attribution of menus to hole: We determined the spatial preference for each rat by establishing the percentage of activation of each hole during the last ten sessions. To favor the emergence of Anti-fragile choices, menus' attribution was made as follow:

   - Anti-fragile exposure was associated to the preferred hole
   - Robust exposure was associated to the 2nd preferred hole
   - Vulnerable exposure was associated to the 3rd preferred hole
   - Fragile exposure was associated to the least preferred hole

5. Training on the exposures: Here, animals were first trained on the gain domain, i.e. no time-out punishment, for each menu/hole. They were subjected to two 20-min sessions, with unlimited number of pellets, during which only one hole was available (eight training sessions in total). For Anti-fragile and Vulnerable options, we used the two sequence-types where the jackpot was available at the 10th and 60th activations (Supplemental Table S1) to ensure that all individuals could experience both an early and delayed jackpot during training. Animals were then allowed to explore all gain options (four opened holes) during 9 20-min sessions, for which different behavioral sequence-types were used. Before training on the loss domain of the different menus, animals were first exposed to a mild and constant 3-sec time-out punishment. Here, animals had access to all options, but half of the activations lead to a 3-sec time-out

Figure 7: Time-line of experiments



punishment, notified by a 3-sec tone and the extinction of the house light. During this period, pokes in the different holes or the food magazine had no programmed consequences. After the 3-sec time-out punishment, the house light was turned on and the animals could again pokes in the different holes. Following nine 20-min training sessions, the loss domain of each menus was progressively introduced. As described above, each time-out punishment was notified by a 3-sec tone and the extinction of the house light for the whole duration of the punishment, which termination was signaled by house light illumination. Animals were first exposed to concave exposures, having only access to vulnerable and fragile holes during four 20-min sessions. They were then exposed to convex losses (only Anti-fragile and Robust holes available) for another four 20-min sessions. Thus, at the end of the training, all animals experienced both extreme events at least four times.

6. Final tests: For the final tests, animals were free to explore all menus throughout the forty 20-min sessions. Animals experienced four times the ten different sequences, randomly distributed across sessions. Population (n = 20) was subdivided into two groups that experienced different, but equivalent, sequence-type distribution (Supplemental Figure S2)).

Figure 7 depicts a simplified time-line of the experimental procedure.

Figure S3 shows the chain of events in the ten sequences used for behavioral training and testing. Numbers in the second column indicate which event would occur and the sign preceding it whether it belongs to the gain (no sign) or loss (minus sign) domains. For example, the third event in the first sequence (noted -2) triggers a time-out punishment of 9 or 12 sec, depending on whether an animal performed his third nose poke in concave or convex menu, respectively. Note that extreme events (which should be noted 4 or -4) do not appear in the sequences, as they would automatically replace the 10th or 60th events of the sequence. See supplementary material for detailed information on sequences of events.

Modelling and Statistical Analysis

Modelling Convex and Concave Exposures

Central to our experimental design is the notion of convex/concave exposure under radical uncertainty, that is, when probabilities and consequences are unknown *a priori* to subjects. A well known measure of convexity is the Jensen gap that is derived from Jensen's inequality (see [28] for a graphical exposition), which we now define and relate to statistical moments. In our context, the relevant form of Jensen's equality states, loosely speaking, that the expectation of a convex function of a random variable is larger than the value of that function when evaluated at the expectation of the random variable. Jensen's gap is then defined as the difference between the expectation of the function minus the value of the function at the expectation (hence positive by construction). The inequality is reversed for a concave function and the (positive again) Jensen's gap is then defined as the difference between the value of the function at the expectation minus the expectation of the function. In the gain domain, rats obtain 1, 3, 12 or 80 sugar pellets if the convex exposure is chosen, or 2, 4, 5, 5 pellets if the concave exposure chosen. In the loss domain, convex exposure imposes 6, 12, 15 or 15 seconds of time-out punishment, and 3, 9, 36, 240 seconds for the concave exposure. Because the values in the loss domain are proportional to the values in the gain domain, the former corresponding to 3 times the latter, we focus here on values for pellets. The statistical properties of convex and concave losses follow accordingly. More formally, sequences of gains are ordered by increasing values and are denoted $\{x_i^{\mathcal{X}}\}_{i=1}^4 = \{1, 3, 12, 80\}$ for the convex exposure and $\{x_i^{\mathcal{V}}\}_{i=1}^4 = \{2, 4, 5, 5\}$ for concave gains. We assume identical probabilities $\{p_i\}_{i=1}^4 = \{p_1, p_2, p_3, \varepsilon\}$ for both concave and convex gains, where ε is the ex-ante probability of the rare and extreme event (REE for short). The third value (that is, 12 for convex gains and 5 for concave gains) is labeled a rare event (RE), and the sets of the lowest two values (that is, 1 and 3 for convex gains, 2 and 4 for concave gains) for both exposures are composed of normal events (NE).

Making use of the exponential transform, we define Jensen gaps for the convex and concave exposures as, respectively, $\mathcal{JG}_{\mathcal{X}} = \sum_{i=1}^4 p_i e^{x_i^{\mathcal{X}}} - e^{\sum_{i=1}^4 p_i x_i^{\mathcal{X}}}$, $\mathcal{JG}_{\mathcal{V}} = \sum_{i=1}^4 p_i e^{x_i^{\mathcal{V}}} - e^{\sum_{i=1}^4 p_i x_i^{\mathcal{V}}}$, which relate to statistical moments as follows. The moment generating function for the convex and concave exposures are given by, respectively: $M_{\mathcal{X}}(t) = \sum_{i=1}^4 p_i e^{t x_i^{\mathcal{X}}}$ and $M_{\mathcal{V}}(t) = \sum_{i=1}^4 p_i e^{t x_i^{\mathcal{V}}}$. The series expansion of the exponential function allows us to derive:

$$M_{\mathcal{X}}(1) = 1 + m_1^{\mathcal{X}} + \frac{1}{2!} m_2^{\mathcal{X}} + \frac{1}{3!} m_3^{\mathcal{X}} + \cdots + \frac{1}{N!} m_N^{\mathcal{X}} + \ldots,$$

$$M_{\mathcal{V}}(1) = 1 + m_1^{\mathcal{V}} + \frac{1}{2!}m_2^{\mathcal{V}} + \frac{1}{3!}m_3^{\mathcal{V}} + \cdots + \frac{1}{N!}m_N^{\mathcal{V}} + \ldots$$

where $m_N^{\mathcal{X}}$ and $m_N^{\mathcal{V}}$ are the N-th statistical raw moments of the convex and concave exposures, respectively. For example, $m_N^{\mathcal{X}} = \sum_{i=1}^{4} p_i (x_i^{\mathcal{X}})^N$ for any integer $N \geq 1$. Jensen's gaps for both exposures are then given by $\mathcal{JG}_{\mathcal{X}} = M_{\mathcal{X}}(1) - e^{m_1^{\mathcal{X}}}$, $\mathcal{JG}_{\mathcal{V}} = M_{\mathcal{V}}(1) - e^{m_1^{\mathcal{V}}}$. Using again the series expansion of the exponential of the first moment for each exposure, it follows that both positive Jensen gaps can be written in terms of moments, as follows:

$$\mathcal{JG}_{\mathcal{X}} = \sum_{k=2}^{\infty} \frac{m_k^{\mathcal{X}} - (m_1^{\mathcal{X}})^k}{k!}, \quad \mathcal{JG}_{\mathcal{V}} = \sum_{k=2}^{\infty} \frac{m_k^{\mathcal{V}} - (m_1^{\mathcal{V}})^k}{k!} \tag{3}$$

Note that the sums in equation (3) nicely allow for a straightforward decomposition in terms on all raw moments with order larger than two.

More specifically, we set probabilities to $\{p_i\}_{i=1}^4 = \{0.5 - \varepsilon/3, 0.4 - \varepsilon/3, 0.1 - \varepsilon/3, \varepsilon\}$ identically for both concave and convex gains. Treating the probability for the REE as a varying parameter, it can be shown that Jensen's gaps for convex and concave exposures are ordered, with the former larger than the latter when $\varepsilon \leq 0.02$. Over that range, it also holds true that both Jensen's gaps are monotone increasing function of ε. While this property also holds for the expectations and variances of both exposures, it doesn't for higher-order moments. For example, the skewness and kurtosis of the convex exposure turn out to be hump-shaped functions of ε, with peaks corresponding to values of ε smaller than one percent. This fact underlines that convexity measured by Jensen's gap offers a unifying approach to rank exposures, seen as lotteries, whereas statistical moments do not necessarily do so. We conjecture that, more generally, all lotteries satisfying our assumptions (including monotone probability distribution) can be ranked according to their convexity as measured by the Jensen gap.

In Table 2, we report the central moments and the log of Jensen gaps in the top five rows, for both exposures when $\varepsilon = 1/75 \approx 1.3\%$, which corresponds to obtaining one REE out of 75 nose pokes. For that particular parametrization, convex and concave exposures differ mostly in their respective kurtosis and in their Jensen gaps. From row six and below, we report the decomposition of Jensen gaps in terms of the raw moments. For example, in row six is given the ratio of $(m_2 - (m_1)^2)/2$ - that is, half the variance - to the corresponding Jensen gap for each exposure, in percentage. Strikingly, while moments of order 2 to 8 explain about 90% of the concave exposure's Jensen gap, they contribute negligibly to that of the convex exposure. In fact, it turns out that moments with order around 80 start contributing, although for a small share each, to the convex exposure's Jensen gap. In sum,

Table 2: Statistical moments and decomposition of Jensen gaps for convex and concave exposures

	Convex Exposure	Concave Exposure
Mean	3.89	3.11
Std Deviation	9.37	1.14
Skewness	7.16	0.20
Kurtosis	57.85	1.38
$\log(\mathcal{JG})$	75.68	2.93
m_2 (%)	10^{-32}	3.48
m_3 (%)	10^{-30}	11.14
m_4 (%)	10^{-29}	18.31
m_5 (%)	10^{-28}	20.65
m_6 (%)	10^{-27}	17.96
m_7 (%)	10^{-26}	12.84
m_8 (%)	10^{-25}	7.86
m_{80} (%)	4.45	10^{-66}

while for the concave exposure a few low-order moments concentrate the contributions to the Jensen gap, the contributions of raw moments to the convex exposure's Jensen gap spread across a larger number of very high order raw moments.

We derive next the ex-ante properties of the probability distributions associated with concave and convex gains (for definitions of stochastic dominance and congruent utility classes, see for example Fishburn and Vickson [15]; also note that first-order, resp. second-order, stochastic dominance implies second-order, resp. third-order, stochastic dominance):

- In the domain restricted to NE, concave gains first-order stochastically dominate convex gains. In addition, concave gains then have a larger expected value than that of convex gains, with equal variance, skewness and kurtosis for concave and convex exposures.

- In the domain restricted to NE and RE, concave gains second-order stochastically dominate convex gains. In addition, concave gains then have a larger expected value and smaller variance, skewness and kurtosis than that of convex gains.

- In the full domain including NE, RE and REE, convex gains second-order stochastically dominate concave gains if and only if $\varepsilon \geq 0.302\%$. In addition, convex gains then have a larger expected value, variance, skewness and kurtosis than that of concave gains if $\varepsilon \geq 0.302\%$.

The above assumptions on the experimental design are stated in terms of stochastic dominance and moments of the probability distributions. They also have implications in relation to standard approaches to decision-making. First, from an ex-ante perspective with perfect information about the probability distributions of all exposures, value-maximizing subjects whose preferences are represented by any non-decreasing value function would choose concave gains in the domain restricted to NE. They would continue to do so in the domain restricted to NE and RE for any non-decreasing and concave value function. However, in the full domain with REE, value-maximizing subjects endowed with any non-decreasing and concave value function are predicted to choose convex gains.
We denote $\mathbb{V}(\{x_i\}_{i=1}^n, \{p_i\}_{i=1}^n)$ the value attributed to sequences of gains $\{x_i\}_{i=1}^n$ with probabilities $\{p_i\}_{i=1}^n$, for $n \leq 4$. The above assumptions have the following implications in terms of value maximization:

- In the domain restricted to NE, $\mathbb{V}(\{x_i^{\mathcal{V}}\}_{i=1}^2, \{p_i\}_{i=1}^2) > \mathbb{V}(\{x_i^{\mathcal{X}}\}_{i=1}^2, \{p_i\}_{i=1}^2)$ holds.
- In the domain restricted to NE and RE, $\mathbb{V}(\{x_i^{\mathcal{V}}\}_{i=1}^3, \{p_i\}_{i=1}^3) > \mathbb{V}(\{x_i^{\mathcal{X}}\}_{i=1}^3, \{p_i\}_{i=1}^3)$ holds.
- In the unrestricted domain, $\mathbb{V}(\{x_i^{\mathcal{X}}\}_{i=1}^4, \{p_i\}_{i=1}^4) > \mathbb{V}(\{x_i^{\mathcal{V}}\}_{i=1}^4, \{p_i\}_{i=1}^4)$ holds if and only if $\varepsilon \geq 0.302\%$.

Note that the reversal in second-order stochastic dominance (and value) in the full domain favors convex gains when extreme events that are indeed very rare - with probabilities much smaller that one percent - are added. In contrast, absent REE, adding only a RE is not enough to favor convex gains, even though it allows subjects to possibly detect convexity/acceleration and concavity/deceleration of gains and losses.

The above assumptions hold true under expected utility, that is when $\mathbb{V}(\cdot, \cdot) = \sum_{i=1}^n p_i u(x_i)$ with the appropriate conditions on the utility function u (non-decreasing for first-order stochastic dominance, non-decreasing and concave for second-order stochastic dominance; see [15]). To better match experimental data, however, expected utility is increasingly

supplemented by some form of probability weighting $w(p_i)$. For instance, [10] use a two-parameter functional form due to [35] and find that 30 rats out of 36 behave as if their probability weighting $w(p_i)$ is concave. In that case, the above ex-ante properties hold *mutatis mutandis*, for example, under expected utility with probability weighting, defined as $\sum_{i=1}^{n} w(p_i)u(x_i)$. Alternatively, in the setting of rank-dependent expected utility, we can make use of results 3 and 4 in [24] to show that the above ex-ante properties still hold provided that the transformation of the cumulative distribution function is any increasing-concave function. Note that since in our experimental design REEs are both extremes (in the sense of being the largest values) and rare (that is, they have very low probabilities), one expects similar results under probability weighting and rank-dependent expected utility (or cumulative prospect theory for that matter). Although in theory assuming simply a weighting function $w(p_i)$ that is inverse-s-shaped and "very" convex for large gains (see [35] for the associated parametric restrictions) could overturn the rankings of concave and convex exposures stated above, results in [10] suggest that this is not to be expected for the overwhelming majority of the rats that are subject to their experiments.

Choice Data Analysis

Given the four options modelled in the previous section, each of the 20 rats went through 41 final sessions. Denote f_A, f_F, f_R, and f_V the relative frequencies (in terms of nose pokes) with which each of the four options Antifragile, Fragile, Robust and Vulnerable, respectively, is chosen over a particular session. In order to represent the rats' choices, we construct the rotated square in panel (d) of Figure 1 using a linear transformation of the 3-frequency vector, as follows:

$$\begin{pmatrix} \text{TSREE} \\ \text{OSREE} \\ 1 - f_F \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} f_A \\ f_R \\ f_V \end{pmatrix} \tag{4}$$

In equation (4), the first two rows deliver the two coordinates of interest that are the Total Sentivity to REE (TSREE in short) and One-sided Sentivity to REE (OSREE), while the last row requires that all frequencies sum to one. To derive TSREE and OSREE, it is useful to go through the following steps. Since we are interested in the convexity property of each chosen option, we first measure the frequencies of convex options in the gain domain by summing f_A and f_V. Similarly, $f_A + f_R$ measures how frequent choices are convex in the loss domain. It follows that TSREE measures total convexity if it equals $2f_A + f_V + f_R$, while OSREE equals $f_V - f_R$ and measures asymmetric convexity (that in the gain domain

minus that in the loss domain). Hence the first two rows in equation (4). Using that $f_A + f_F + f_R + f_V = 1$ - that is, the last row of equation (4) - it follows that TSREE equals the simpler expression $1 + f_A - f_F$. Note that the 3×3 matrix in equation (4) is invertible, which implies that all three frequencies can be recovered from the left vector. In addition, that matrix is not unique in the sense that the definitions of TSREE and OSREE that appear above as its first two lines could be moved to any other rows, provided that the identity in the third row and the constraint in the last row are adapted accordingly. This essentially means that given both differences TSREE ($= 1 + f_A - f_F$) and OSREE ($= f_V - f_R$), any of the four relative frequencies provides enough information to derive the remaining three, using the constraint that all four frequencies sum up to one. Finally, it follows that in the rotated square in panel (d) of Figure 1, the four vertices have the following coordinates: $(0, 0)$ for Fragile, $(0, 2)$ for Antifragile, $(-1, 1)$ for Robust, and $(1, 1)$ for Vulnerable. TSREE and OSREE are depicted in Figure 2.

TSREE and OSREE are also related to behavioral measures of the share of nosepokes that lead to either exposure or non exposure to REE, as follows: we define Black Swan Avoidance as the share of nosepokes that lead to avoidance of Black Swans - i.e. $f_A + f_R$ - and Jackpot Seeking as the share of nosepokes that lead to be exposed to Jackpots - i.e. $f_A + f_V$. It follows that TSREE is by definition the sum of those two shares, while OSREE is the difference between the former and the latter. Formally, if we denote JPS and BSA the shares of nosepokes that lead to be exposed to Jackpots and to avoid Black Swans, this implies that JPS$=$ 0.5(TSREE$+$OSREE) and BSA$=$ 0.5(TSREE$-$OSREE). In other words, in the (not rotated) square that appears in panel (d) of Figure 1, the four vertices have the alternative coordinates: $(0, 0)$ for Fragile, $(1, 1)$ for Antifragile, $(0, 1)$ for Robust, and $(1, 0)$ for Vulnerable, in the (JPS,BSA) coordinates. Obviously, both interpretations are equivalent up to a change in coordinates which is a bijection. Note that in the rotated square in panel (d) of Figure 1, lines with a 45-degree slope depict choices with constant BSA so that lines moving north-west represent increasing BSA. Similarly, lines with a -45-degree slope depict choices with constant JPS so that lines moving north-east represent increasing JPS. Median BSA and JP are depicted in Figure 3.

In Figure 4 we report the short term behavioral responsiveness of rat to REE. Behavioral responsiveness is measured by calculating for each rat before/after differences in TSREE and OSREE. This amounts to calculate TSREE and OSREE over the 10 choices preceding a REE and over the 10 choices following the occurrence of each REE. The window of 10

Table 3: Before/after differences for Total Sentivity to REE and for One-sided Sensitivity to REE, averaged over the 41 sessions, for each of the 20 rats - see Figure 4. We highlight negative mean differences in red and positive mean differences in yellow.

	Total Sensitivity to REE		One-sided Sensitivity to REE	
Rat	jackpot	black swan	jackpot	black swan
• 1	-0.062	-0.133	-0.046	-0.467
• 2	-0.025	-0.108	-0.075	0.215
• 3	0.050	0.050	-0.100	-0.083
• 4	-0.036	0.100	-0.036	-0.300
• 5	0.010	0.400	0.010	-0.067
• 6	0.100	0.117	-0.100	0.017
• 7	-0.650	0.186	0.050	0.000
• 8	0.017	0.111	-0.050	-0.044
• 9	-0.264	-0.210	0.009	-0.070
• 10	-0.045	0.400	-0.009	0.200
• 11	-0.175	0.080	0.042	-0.027
• 12	0.033	0.150	-0.122	0.000
• 13	0.085	0.150	0.054	-0.350
• 14	-0.114	-0.100	-0.086	-0.500
• 15	0.167	0.317	-0.011	-0.050
• 16	-0.000	-0.100	-0.062	-0.100
• 17	0.300	0.150	0.100	-0.107
• 18	-0.190	0.600	-0.010	0.000
• 19	0.105	0.160	-0.053	-0.040
• 20	0.014	-0.100	-0.029	-0.300

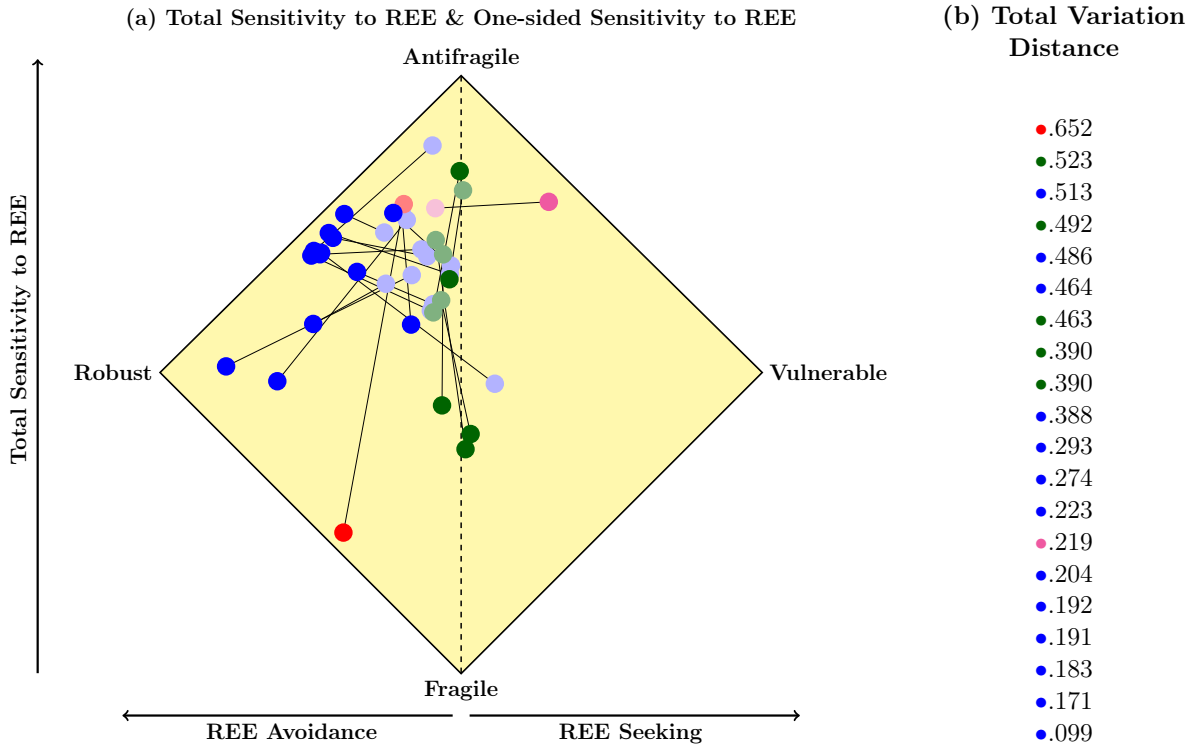
choices before and after is dictated by the configuration of REE in our design. A REE can happen at 11th choice in a sequence (see appendix). Before/after TSREE and OSREE are averaged by rat over the 41 sessions in the gain domain (Jackpot) and in the loss domain (Black Swan). Table 3 reports the values for the mean before/after differences. In each panel of Figure 4, each rat is depicted by a color-coded dot and the dotted black line represents the 45-degree line. Color-coded dot on the 45-degree line indicate no difference in TSREE or OSREE before and after a REE, i.e. no short term behavioral responsiveness to REE. To visualisation purpose, we computed for each panel a smooth spline regression that estimates a non parametric relationship between before and after TSREE and OSREE. They are plotted as solid black lines in each panel.

We then test the statistical significance of short term behavioral by conducting bootstrap paired-sample mean tests on before/after coordinates (separately) for each type of REE. An observation in the sample is the mean behavioral responsiveness in TSREE or OSREE for a given rat. Observations therefore respect statistical independence. Bootstrap tests lead to the following p-values, under the null hypothesis that the mean difference is zero. Following Jackpots, $p = 0.3470$ for TSREE and $p = 0.0522$ for OSREE, which indicates that the null hypothesis is not rejected at 1%. Following Black Swans, however, the null hypothesis is rejected at 1% since $p = 0.0045$ for TSREE and $p = 0.0064$ for OSREE. In sum, the mean before/after difference following a REE in the loss domain is significant for both Total and One-sided sensitivities, further confirming the Black Swan avoidance result that we document in this paper.

We also report in Figure 8 how choices by rats after the 41 final sessions compare to choices made in the training step 3, defined in the section on Behavioral Training and Testing. We do so by computing TSREE and OSREE in step 3 sessions and pair coordinates resulting from training with that resulting from the 40 experimental sessions reported in Figure 2. The lines between two color-coded dot connects for each rat its conditioning coordinates with its coordinates resulting from its behavior in the 41 sessions. We report on the right side of the figure the total variation distance between the distribution of nosepokes in the 4 holes in the training and the distribution in the 41 experimental sessions. The interpretation of total variation distance in our setting is simple: it measures, for each rat, the proportion of nosepokes that need to be changed in order to equalize the behavior in the training and the behavior in the 41 final experimental sessions. Median and mean total variation distance are 0.293 and 0.324 respectively, which implies that half of the rats changed more than 30% of their choices in the final experimental sessions compared to the training sessions.

Finally, in Figure S1 of the Supplementary Material, we report what we label "convexity premiums", which are counterfactual situations that we construct as follows. Given the rats' choices over the 41 sessions, we recall all random sequences that have been used to generate gains and losses for each rat. We next compute the outcome that would have been obtained, had the rat chosen convex options in the gain domain (that is, Antifragile or Vulnerable) and in the loss domain (that is, Antifragile or Robust) all the time. The right and left panels in Figure S1 depict the normalized convexity premiums thus computed, for each rat represented in rows, with convex exposures yielding more pellets and implying less waiting in terms of seconds.

Figure 8: Variation of choices observed in final experimental sessions compared to that observed in training sessions (step 3 in Material and Methods section on Experimental Method Details); panel (a) shows variations of Total Sensitivity to REE (y-axis) and of One-sided Sensitivity to REE (x-axis), that is, dark color-coded dots replicate averages over the 40 sessions for each of the 20 rats - see panel (c) in Figure 2 - and are paired with averages over training sessions for the same rat, depicted with light color-coded dots connected through lines with corresponding dark color-coded dots; panel (b) shows total variation distance, that is, the proportion of total nose pokes in training sessions that need to be changed to replicate nosepokes in final sessions



Augmented Q-Learning Model Estimation and Simulation

Parameter optimization is carried out for each rat over the 41 sessions, using observed choices and outcomes. We estimate by maximum likelihood two series of augmented Q-Learning models. The first series integrates REE in Q subvalues whereas the second series does not. Each series consists of a first baseline model without specific forgetting rates and decision weights attached to REE. Two partial models integrate either specific forgetting parameters or decisions weights on the presence of REE. The fourth model integrate both.

In total, we estimate eight nested models for each rat. Formally, we estimate four different models: a baseline model (Model 1) without specific forgetting parameters ($\alpha_k^f = \alpha_k$, with $k \in \{g, l\}$) and with zero decision weights of REE ($\gamma_g = \gamma_l = 0$). Model 2 introduces specific forgetting parameters ($\alpha_k^f \neq \alpha_k$, with $k \in \{g, l\}$) and Model 3 introduces decision weights of REE (γ_g and $\gamma_l \neq 0$). Model 4 introduces both.

To ensure convergence and avoid local maxima, each parameter in baseline models was initialized with two different points of its parameter space. This makes in total $4^2 = 16$ combinations of initial parameter values. Each of the 16 models is estimated using an automatized two-step procedure. The fist step consists of an unconstrained downhill simplex method in [29], which does not make use of first derivatives. Estimated parameters are then used as initial parameters in the second step of the estimation procedure, a quasi-Newton method in [16]. Parameter estimates from baseline models are then used as initial parameters for subsequent augmented models. Estimation is carried out using the above two-step procedure to ensure that we convergence is reached.

Once the eight models are estimated, we then compare them using both Akaike information criterion (AIC) and Bayesien Information criterion (BIC), that penalizes the use of additional parameters, so as to select which model configuration best fits observed behavior for each rat (see Supplementary Material for detailed results per rat). We check whether the selected model properly estimates sensitivities to REE by comparing estimated sensitivities to observed sensitivities to REE. This is done by computing Total Sensitivity to REE and One-sided Sensitivity to REE based on estimated choice frequencies from the extended Q-Learning model over the 41 sessions. Fitted sensitivities are reported bottom right panel in Figure 5. Finally, we checked that selected models were able to reproduce experimental data by running simulations for each phenotype observed in the sample of animals (as in [32]). Simulations are conducted over the artificially generated original 41 sessions, using individual parameter estimates for pink and red rats, median parameter estimates for blue and green rats and estimated Q subvalues as prior Qs. For each simulation, we then computed Total Sensitivity to REE and One-sided Sensitivity to REE. Simulations are reported in bottom right panel of Figure 5. See Supplementary Material for more details on the estimation results and for an exploration in related outcome range-adaption models.

Acknowledgments

Funding: The authors are grateful for the support of ANR through BEAM (ANR-15-ORAR-0004-03) and of CNRS through a MITI interdisciplinary grant (call for projects on rare events).

Author contributions: C.B., M.D, S.L., and P.P. designed the experiments. M.D and L.W.M. performed the experiments. S.L. and P.P. analyzed and modelled the data. S.L, P.P. and C.B, wrote the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All data and codes needed to replicate the conclusions are available in the Supplementary Material and upon request from the authors.

Supplementary Material

Sequences used in the experimental design

In Table S1 we report the different types of sequences used for behavioral training and testing, which shows the position of the extreme events. For example, in the sequence-type 6, jackpot could be obtained at the 10th position while the Black-swan would be triggered at the 60th activation.

Table S1: Position of REE in the sequence

#	Jackpot	Black Swan
1		
2	10	
3	60	
4		10
5		60
6	10	60
7	60	10

Table S2 shows the different behavioral sequence-types used for the forty testing sessions. As indicated in the text, half of the population was subjected to the sequence-types described in the left part of the 'type column', while the other half experienced the sequence-types described in the right part of the 'type column'. Table S3 shows the succession of events for each sequence.

Table S2: Sequences and types

Session	Sequence	Type	
1	2	5	3
2	10	6	5
3	5	1	3
4	3	7	1
5	4	3	3
6	1	7	6
7	6	2	6
8	6	6	2
9	4	2	4
10	1	1	1
11	7	7	4
12	1	3	2
13	7	7	7
14	8	2	4
15	5	2	5
16	7	2	6
17	3	1	6
18	5	5	4
19	2	1	6
20	7	6	1
21	9	4	4
22	4	2	6
23	2	5	2
24	2	3	5
25	4	2	2
26	9	2	2
27	10	5	6
28	10	5	4
29	10	1	4
30	3	3	6
31	8	5	4
32	3	7	6
33	6	2	1
34	9	4	4
35	9	3	5
36	6	5	1
37	1	5	3
38	8	4	4
39	5	3	7
40	8	4	4

Table S3: Succession of events for each sequence

#	Succession of events
1	2,-1,-2,-2,2,1,2,-1,2,1,-2,1,2,-1,-3,1,-1,-2,-2,-1,1,1,3,1,1,1,1,-1,1,-2,2,-1,-2,1,-3,-1,2,-1,-2,-1,-1,1,2,-1,-3,1,-3,-1,-2,-1, -2, 1, 2,2,-1,-1,-2,1,2,1,-1,1,2,-3,-2,-1,-1,-1,-1,-1,-3,-1,-2,-1,-2,-1,-2,-1,3,-2,2,2,-1,1,-1,1,-2,1,-2,1,-2,1,2,1,-1,-2,1,1,1,-1,1,2,2,-1, 1,-1,2,-1,1,-1,-2,2,2,1,1,-1,1,1,-2,1,3,1,-1,1,1,1,-2,1,-2,1,-2,-1,2,-1,2,-1,-1,-1,1,-1,1,-1,2,1,-2,1,-2,-1,2,-1,-1,-1,-1,1,-2,1,-1, -1,-3,2,2,2,2,-1,-2,-1,-2,1,3,1,2,1,-1,1,2,-1,-3,1,2,1,-2,-2,1,1,-3,-1,1,1,-2,1,2,2,-3,1,1,-1,3,-2,-2,1,2,-1,2,1,-2,1,-1,2,-1,-1,-3, 2, 2,-2, 2,1,-1,-1,2,-1,-2,1,2,-1,3,1,1,1,-2, 2,-3,1,-2,1,2,1,-1,-1,2,1,3,-1,-3,1,-1,1,1,-1,-1,-2,-3,-1,-3,1,1,1,-1,1,-2,-1,3,-1, 1, 1,-2,2,-3,-1,1,1,2,-1,-2,1,2,1,3,-1,-2,2,1,2,-1,1,-1,1,2,-1,3,-1,-2,-2,-1,-1,-1,-1,1,-1,1,1
2	-2,1,-1,-2,-1,-1,-3,1,2,1,-2,2,2,1,1,2,-2,-1,3,-2,1,1,-1,2,2,1,-3,-1,-2,-1,-2,1,-2,1,-1,1,-2,-2,1,1,-2,-1,-2,-2,-2,1,2,2,1,1,-2,-1, 1,-1,-1,1,-2,1,-2,1,-2,2,2,1,1,1,1,-1,-1,1,-1,1,-2,2,-3,-2,-3,-1,1,1,-2,-1,1,-1,-2,1,-3,1,-1,1,-1,-2,-1,3,-2,-1,1,2,-1, 2, 1,1, 1,3,-1,2,-1,2,-1,-1,-1,3,1,-3,1,2,-1,2,-1,1,1,-1,-1,-2,-1,-3,-1,-1,-1,-2,-1,-2,-1,-2,1,3,1,-2,1,-2,2,-3,-1,-2,2,1,2,-3,1,3,1, -2, 1, 2,2,-1,-1,3,-1,-2,1,-3,-1,2,1,2,1,1,1,-1,1,-1,-1,3,-1,2,-1,-3,-2,3,1,-1,-2,2,1,2,-1,-2,-1,2,1,1,-1,1,-1,-2,-1,1,-1,1,-1,1,-1,1,-2,- 2,1,1,1,-2,1,2,1,1,-1,-2,-1,1,2,-2,-1,2,-1,-2,-2,-2,-1,-2,1,2,-1,2,-2,-1,1,3,-1,-3,1,-1,1,-3,1,-3,-2,3,1,2,-1,3,-1,3,-1,-3,1,-1,-1, 2,2,1,1,-1,-1,-2,-1,-2,1,2,-1,-1,-1,3,2,-2,1,-2,-1,-1,1,-2,2,3,1,-2,1,2,-1,-2,-1,-1,-1,2,-1,-1,-2,-1,-1,-3,1
3	-2,1,-2,-1,-2,-1,-2,-1,-3,-1,1,1,2,-1,2,1,2,1,-3,-1,-2,-1,1,1,2,-2,-2,-1,2,-1,-2,-3,3,-1,1,-1,1,2,1,1,-1,-1,-1,1,-1,-1,-2,-2,-1,-1, 2,1,3,-1,-2,1,-2,1,-1,1,-2,1,2,-1,2,1,-2,1,2,-1,2,-1,1,-1,-2,-1,-3,1,-2,1,-2,1,-2,-1,2,1,2,-1,-2,1,-2,1,-1,-1,-2,-1,3,1,-2,-1,3,-1,-1,-1,-3, -1,-2,2,1,2,1,-1,1,-1,-2,-1,2,-2,-1,1,1,1,-1,-1,-2,-2,3,1,-3,2,2,2,-1,-1,-2,-1,-2,1,-1,1,2,-1,-2,-1,-1,1,1,1,-1,-2,2,1,2,1,-1,1,1, -1,-1,-1,2,1,2,1,2,-1,-1,-1,2,-1,-1,1,2,-2,-2,1,-1,1,1,-2,-1,-1,2,-1,-3,-1,3,-2,-1,1,-1,-1,-1,-1,2,-1,-2,1,-1,-1,-1,-1,-1,1,1,-1,2, 1,1,2,2,-1,-1,1,-1,-2,1,1,-1,1,-1,-2,-1,1,1,2,-2,3,-3,1,-1,-3,-1,1,-2,-1,1,1,-2,1,-1,3,1,-1,-1,-2,-1,2,-2,2,1,3,1,1,1,-3,-2,-1,1,-2,1,3,-1, -1,-1,-2,2,2,-1,2,1,2,-1,2,1,-1,-1,1,-2,1,1,2,-1,2,1,-3,1,3,1,2,-1,-2,-1,-3,1,-2,2,-3,1,1,-1,3,1,-1,2,1,1
4	1,-1,-2,2,-1,1,-2,1,-2,1,-3,-1,-2,2,-1,1,2,1,2,-2,-2,1,-2,-2,2,1,-1,-1,-2,-1,-1,-1,-1,2,-1,-1,-2,2,3,-1,-1,1,-2,-1,2,-1,1,-1,-2,1,- 2,-1,1,-1,2,1,2,1,2,-1,-3,1,-1,1,-2,-1,-3,-1,1,-1,2,2,-3,-1,1,-2,-1,-1,-1,1,2,-1,2,-1,-3,-1,-1,-1,-2,1,3,1,-2,-2,1,-2,2,1,-3,-1,1,- 1,-3,1,-2,1,2,-1,-1,-1,1,-1,-3,1,-2,-2,-3,-2,-2,1,-2,-1,-2,-1,-2,1,1,1,-2,1,2,1,-2,1,2,-1,-1,1,-1,-1,-3,1,-2,-1,-2,-1,-2,-1,1,1,-2, 1,1,-2,2,-1,-1,1,-3,-2,-3,1,-2,1,2,-1,2,2,1,2,-1,-1,3,2,-2,-1,2,-2,-2,-1,2,-2,-1,1,2,-1,1,-1,1,-1,-2,-1,-2,1,3,-2,-3,1,-3,-1,1,1,3,1,- 2,-1,-1,-2,-2,1,2,-2,-2,-1,1,1,2,-1,-2,1,3,1,2,1,-1,-1,-2,1,1,-2,1,1,-1,-1,-1,-1,2,1,-2,1,1,-1,2,1,2,-1,1,2,-2,1,3,1,1, -2,-2,1,- 1,2,1,-1,-2,-1,-2,1,-2,-1,-3,2,-2,1,-1,-1,-2,-1,-3,2,-2,-1,1,-1,2,-2,-3,-1,3,-2,1,1,-1,-1,1,1,2,1,2,-1,2,1,-1,1
5	1,-1,2,2,3,-1,2,-1,2,1,-2,1,2,-1,-1,-1,-2,-1,-1,2,-3,-1,-1,3,2,1,2,-1,1,1,-2,-1,2,-2,1,1,-2,-1,-2,1,-2,-1,-2,1,1,-1,-3,1,1,-1,-1,-2,2,- 1,-2,-2,-2,1,-1,1,-1,-1,2,-1,2,1,-1,2,-2,1,2,-1,2,1,2,-2,-2,1,1,1,-2,-1,-2,-1,-2,-2,-1,2,3,1,1,-1,-2,1,1,-1,2,1,2,1,-3,1,2,-1,2,-1,- 2,1,2,1,-1,-2,-1,-1,2,1,2,2,-2,1,2,-1,-2,-1,-3,-1,-1,-1,-1,1,2,-1,-1,-3,2,1,-1,-1,3,1,-1,1,1,-1,3,-1,-2,1,-1,-1,3,1,2,1,-1,1,2,-1,2,1,- 3,1,-2,1,-2,1,2,1,-1,-2,-1,-2,-2,-1,2,1,2,1,-1,-1,2,-2,2,-1,-3,1,1,1,3,-2,-3,-1,3,1,-1,1,2,1,1,1,1,2,2,1,2,-1,2,1,-3,-1,2,-1,1,-1,- 2,1,1,1,-1,-2,-1,2,-1,-1,-2,1,2,1,2,-1,2,-1,2,-2,-1,-1,-1,-2,1,1,-2,-1,-2,1,-2,1,-1,2,3,1,1,-1,-2,-2,-2,1,-2,1,2,-1,3,2,-2,2,-2,1,1,1,- 2,-2,3,1,-2,-1,-3,-1,3,-1,-2,-1,2,-2,-3,1,-3,-1,1,1,1,1,1,-2,-1,-3,-1,-1,-2,3,1,1,-1

| 6 | -1,1,1,-1,-2,1,-2,-1,-2,2,-1,1,-2,-1,2,1,1,1,-2,-1,1,1,2,-2,2,1,3,-1,2,-1,-3,-1,-2,1,-2,-1,-2,-1,-1,1,-1,1,-3,2,-1,-2,1,1,-1,1,- 2,1,1,1,1,1,1,2,-2,1,-2,2,-2,-1,1,1,-1,-1,-1,1,3,2,-2,-1,3,-1,-1,-1,1,1,-1,-2,2,1,-1,-1,-3,-1,-2,-1,-2,1,1,-2,-2,1,-2,1,2,1,3,1,1,- 1,1,-1,-1,-1,2,-2,3,-2,2,-3,2,1,-1,2,-3,1,3,1,-1,1,-2,-1,1,-1,-1,-1,2,2,-2,-2,1,-1,1,1,-1,2,2,1,-1,1,1,1,-1,2,2,1,-1,1,1,-2,1,3,-2,1,1,3,-1,-3,-1,1,1,- 3,1,3,1,1,-1,-2,-1,2,-1,-1,1,2,-1,1,-1,-1,-1,-2,-1,-2,-1,-3,-1,2,1,-1,-1,2,2,-2,-1,-2,-1,3,1,-2,-1,-1,1,3,-2,-2,-2,-2,-1,-2,2,-2,-1,1,- 1,-2,1,1,-1,-3,-1,-2,1,3,1,-2,-1,-3,1,1,1,-2,-1,1,-1,-2,-1,-1,-1,-1,-1,2,1,1,-2,1,1,1,-2,-1,-1,-2,1,2,-2,3,-1,-3,-1,1,1,-2,1,3,2,1,1,- 2,-2,-2,1,2,1,2,-1,-1,-1,2,1,-1,1,-1,1,-1,-1,-2,-1,1,1,-2,-1,-1,2,-2,-2,1,1,-2,1,2,1,-2,1,2,1,2,1 |
| 7 | -2,-1,-2,1,2,1,-2,-2,1,-1,-1,-1,-3,1,-1,-1,-1,1,2,1,2,-1,-2,-1,-1,1,-2,1,2,1,-2,-1,3,-1,2,-1,-3,1,-2,-1,-2,2,3,-1,-3,1,-2,2,1,1,2,1,- 2,-1,-2,1,3,-1,-1,1,-1,1,3,-3,-3,-1,2,-1,-1,-2,2,1,1,-1,-2,-1,1,1,-2,-1,-2,-1,2,-2,-1,-1,2,-1,2,1,-1,1,2,1,2,1,-2,-2,-1,-1,-1,1,-2,- 1,2,1,-2,-2,-1,1,2,2,-2,-1,2,1,2,1,2,1,-2,1,-1,1,2,-1,2,-1,3,1,-2,1,2,1,-2,-1,-2,1,-3,-2,-3,1,1,-1,-2,1,-2,-2,-3,2,2,2,-2,-1,1,1,-2,1,- 1,-1,-2,1,-3,1,3,-1,-3,-1,2,1,-3,-2,-2,1,2,1,2,-1,-2,1,-3,1,-2,1,3,1,2,2,2,-2,2,-1,-2,1,3,-1,1,1,-2,-1,3,1,-1,1,-3,1,-2,-1,-3,-2,-1,- 2,2,1,1,1,-3,-1,2,-1,-1,2,-1,1,-2,-1,1,1,-1,1,2,-1,3,1,1,-1,2,2,-2,-1,-2,1,1,-2,-2,2,2,-1,1,-1,-2,-1,-2,1,1,-2,-2,1,2,-1,2,1,1,-2,1,- 1,-2,1,-1,-1,1,-2,-2,-1,-2,1,2,1,2,1,1,-1,-1,2,-1,-1,2,3,-1,-1,2,1,2,-1,1,-1,-1,-3,1,2,-2

| 8 | -1,1,-2,-1,-2,-1,-1,1,-2,-1,2,-1,3,1,-2,-1,-2,1,-3,-1,2,-1,3,1,-2,1,-1,1,-2,-1,2,2,-1,-1,-2,1,2,-1,-3,1,-2,1,1,-1,2,1,1,2,-3,-2,2,1,- 3,-1,-1,1,1,1,2,1,-1,1,2,-1,-1,1,2,-1,3,2,-1,-1,-1,1,3,-1,-2,1,-2,-1,-1,-2,2,-1,1,-1,1,-3,-1,2,2,1,2,1,-2,1,-1,1,-1,1,2,1,3,1,3,-1,- 1,-1,-3,1,-2,-1,-1,1,1,-1,2,1,-2,-1,-1,1,-2,1,-1,1,2,1,-2,-1,-2,1,-2,-1,-1,-1,2,1,-1,-1,2,2,2,1,-2,2,1,1,-3,-1,1,-1,-1,-1,1,-1,-1,- 1,2,-1,-2,1,1,-1,2,1,-2,1,1,1,1,-1,1,-1,3,2,-1,-1,-2,1,-1,-2,1,1,1,1,-3,1,-1,2,-2,1,-2,-2,3,-1,-3,1,1,1,-2,1,-2,1,-1,-1,1,1,-2,1,-3,- 1,2,1,-1,1,2,-1,-3,1,2,-1,-2,2,2,-2,-1,1,-1,1,1,1,-2,1,2,-1,-1,1,2,-1,2,-2,-2,-1,-2,1,-2,1,2,2,-2,-1,-3,-1,2,1,1,-1,1,-3,-1,-1,-1,1,- 1,-1,3,1,-2,-2,-1,1,-2,1,-1,-2,2,1,2,1,-3,-1,-3,1,-2,-1,3,-1,3,-2,-3,1,-2,-1,1,-1,-1,1,1,1

| 9 | 1,1,2,-1,-2,1,1,-1,-2,1,1,1,1,-1,-1,-1,-2,-1,-2,1,2,-2,-2,1,1,-2,1,-1,-3,2,-1,1,3,-2,-3,-1,2,1,-2,1,3,2,-2,1,-2,1,-3,1,2,1,-1,1,2,- 1,3,-1,-2,1,2,-2,-1,-1,3,1,2,-1,1,1,-1,-1,1,-1,-2,2,3,1,-2,-1,2,-2,2,1,-1,-1,2,-1,2,-1,2,1,-2,-1,3,-1,-2,2,2,2,1,1,1,2,3,-2,3,-1,-1,- 1,2,-1,1,-1,1,1,-1,-1,-3,1,-1,2,-1,-1,-3,1,-2,-1,-1,1,-2,-2,2,1,-2,-1,-2,2,-2,1,2,-2,1,1,-1,-1,-2,1,1,-2,-1,2,-3,-1,2,-1,-2,-1,2,-1,- 3,2,-1,1,2,-1,1,1,2,-1,-2,1,1,2,2,1,3,-1,1,1,-2,1,2,1,-1,-1,-2,-1,3,1,1,1,1,1,1,-1,-2,-1,-2,2,-3,1,1,-1,3,-2,3,-1,1,-1,-2,1,- 2,1,-1,-1,-2,1,-2,-2,2,-1,-1,1,-2,1,1,-1,1,2,1,-1,-1,1,-3,-1,-2,1,-2,1,-1,-1,-1,-1,1,1,-1,-2,1,-2,-2,1,-1,2,-3,1,-1,-2,1,- 2,1,2,-1,-1,1,-2,-1,3,1,-3,1,2,1,1,1,-2,-1,-1,1,-3,-1,-1,-3,2,3,-1,2,-2,3,-1,2,-1,2,-1,-2,-1,1,-1,1,1,1

| 10 | -1,1,3,-1,-2,-1,-2,-1,3,1,2,1,2,-1,-2,1,-3,1,2,-1,-3,-2,2,-1,3,-2,2,-1,2,1,-2,-1,-2,1,3,2,2,-1,2,1,3,1,1,1,-1,-1,-1,1,-3,1,1,1,2,1,1,- 2,1,2,-1,-1,-2,-1,1,-1,2,1,-2,1,-2,1,-1,1,2,-2,-2,-1,-2,-1,-2,-1,-2,-1,2,-1,-1,-2,-1,-1,-3,1,-3,1,-1,2,-2,1,2,-1,-1,-1,1,1,-2,2,3,1,- 3,1,2,1,-1,1,2,-2,-1,-1,-2,1,1,-1,-1,1,-3,1,-3,2,-3,1,-1,1,-2,1,-2,1,2,-1,3,-1,2,-2,1,-1,2,2,-3,1,-2,2,2,-1,2,2,-2,-1,1,-1,2,1,-1,1,- 1,-1,-1,-1,2,1,1,2,1,1,3,1,3,1,-2,-2,-1,-1,-2,-1,-2,-1,3,-2,1,-1,-1,-1,-1,1,2,-1,2,1,3,-1,2,-2,2,-1,-3,-1,1,-2,1,-1,-3,1,1,-1,-2,1,1,- 1,-2,1,1,1,-1,-1,2,-2,-2,1,2,-2,3,1,-1,3,-3,1,2,-1,1,-1,2,1,2,2,2,-2,-3,-1,-2,-1,2,1,-2,1,-2,1,-1,-1,-1,-1,3,1,-2,-1,2,1,-1,-2,3,-1,1,- 2,2,-2,2,1,-1,2,-2,1,2,1,3,-1,1,-1,2,1,-2,-1,1,-2,1,-1,-1,-1,3,1,2,1,3,1,-3,-1,-1,1

Behavioral measures for each rat

All behavioral measures for each rat that are used in the main text are gathered in Table S4: Total and One-sided Sensitivities to REE, Black Swan Avoidance and Jackpot Seeking (both in %).

Table S4: Behavioral measures for each rat - Total and One-sided Sensitivities to REE - TSREE and OSREE in short; Black Swan Avoidance and Jackpot Seeking in percentage - BSA and JPS in short

Rat	TSREE	OSREE	BSA	JPS
• 1	1.4136	-0.4583	93.59	47.76
• 2	1.2177	-0.1392	67.84	53.92
• 3	0.9749	-0.5808	77.78	19.70
• 4	1.2264	-0.4259	82.61	40.02
• 5	1.4208	-0.4990	95.99	46.09
• 6	1.0819	-0.6024	84.21	23.97
• 7	0.4632	0.0122	22.55	23.77
• 8	1.3706	-0.0669	71.87	65.18
• 9	0.6928	0.0304	33.12	36.16
• 10	1.4652	-0.4058	93.55	52.97
• 11	0.9074	0.0104	44.85	45.89
• 12	1.4365	-0.2772	85.68	57.96
• 13	1.5444	-0.1653	85.48	68.95
• 14	1.5504	-0.3937	97.20	57.83
• 15	1.6610	0.0447	80.81	85.28
• 16	1.4076	-0.5050	95.63	45.13
• 17	0.5661	-0.3929	47.95	08.66
• 18	1.1406	-0.4850	81.28	32.78
• 19	1.6009	0.2997	65.06	95.03
• 20	1.4832	-0.4491	96.61	51.70

Augmented Q-Learning model selection according to BIC

Table S5 presents BIC values for all rat models. In the first column, the rats' phenotypes are color-coded as in the text. Each uniquely selected model is identified by light blue color.

Table S5: BIC values for all augmented Q-Learning models, with selected value highlighted

	REE in Q-subvalues				REE not in Q-subvalues			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
• 1	10147.618	9473.977	9455.796	9233.393	9639.117	8909.843	9048.188	8883.051
• 2	13215.452	13188.041	13121.341	13135.077	13061.968	13040.253	13037.390	13042.841
• 3	8825.413	8748.519	8832.939	8754.428	8668.958	8581.156	8679.882	8592.356
• 4	9559.508	9569.403	9118.611	9127.721	9818.665	9473.840	9338.137	9355.298
• 5	8851.442	8837.212	8737.694	8753.888	8842.836	8831.712	8728.778	8743.304
• 6	7206.668	7026.401	7058.036	7040.684	7113.624	6824.310	6931.265	6812.231
• 7	10194.305	9407.034	9817.474	9419.479	9659.679	9184.589	9559.764	9135.728
• 8	12495.756	12506.051	12505.885	12515.155	12492.775	12496.156	12490.187	12489.188
• 9	12613.911	12183.914	12581.295	12187.231	12194.450	11855.559	12195.525	11831.083
• 10	8822.488	8532.439	8612.817	8624.451	8269.455	8130.544	8083.416	8018.394
• 11	11119.104	11005.450	11006.259	10950.474	11030.692	10873.231	10981.489	10889.932
• 12	11387.450	11198.078	11105.750	11069.321	11222.624	11032.712	11089.737	11065.547
• 13	10835.134	10296.487	10388.582	10241.068	10677.363	10215.775	10373.480	10202.434
• 14	10210.664	9855.573	9536.958	9552.319	9996.584	9880.569	9521.883	9305.535
• 15	9042.715	8933.762	9051.019	8895.580	8964.300	8840.706	8981.304	8824.601
• 16	10395.829	10219.158	9990.742	10003.734	10600.320	10107.624	10019.103	10035.935
• 17	9636.231	9191.330	9164.951	9152.413	9661.941	9285.094	9126.409	9143.220
• 18	12948.075	12737.650	12777.029	12685.432	12808.119	12537.832	12676.942	12523.725
• 19	11146.154	10803.117	10362.629	10371.896	11347.877	10585.857	10230.564	10231.779
• 20	9885.053	9359.053	9279.145	9293.579	9742.212	9304.446	9292.413	9307.507

Augmented Q-Learning model selection according to AIC

Table S6 presents AIC values for all rat models. In the first column, the rats' phenotypes are color-coded as in the text. Each uniquely selected model is identified by light blue color.

Table S6: AIC values for all augmented Q-Learning models, with selected value highlighted

	REE in Q-subvalues				REE not in Q-subvalues			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
• 1	10120.983	9434.024	9415.843	9180.122	9612.482	8869.889	9008.235	8829.780
• 2	13188.983	13148.337	13081.638	13082.139	13035.499	13000.550	12997.686	12989.903
• 3	8799.339	8709.409	8793.829	8702.281	8642.884	8542.045	8640.772	8540.209
• 4	9533.087	9529.770	9078.979	9074.878	9792.244	9434.208	9298.504	9302.455
• 5	8825.399	8798.149	8698.630	8701.804	8816.794	8792.648	8689.714	8691.220
• 6	7180.386	6986.979	7018.614	6988.121	7087.343	6784.888	6891.842	6759.669
• 7	10167.856	9367.360	9777.800	9366.581	9633.230	9144.915	9520.090	9082.829
• 8	12469.528	12466.708	12466.542	12462.697	12466.546	12456.813	12450.844	12436.730
• 9	12587.291	12143.983	12541.364	12133.989	12167.829	11815.628	12155.594	11777.841
• 10	8796.092	8492.845	8573.222	8571.658	8243.059	8090.949	8043.821	7965.600
• 11	11093.143	10966.508	10967.317	10898.552	11004.731	10834.290	10942.548	10838.011
• 12	11361.251	11158.780	11066.452	11016.924	11196.425	10993.414	11050.439	11013.149
• 13	10808.770	10256.941	10349.036	10188.340	10650.999	10176.229	10333.934	10149.706
• 14	10183.890	9815.412	9496.797	9498.771	9969.810	9840.408	9481.722	9251.987
• 15	9016.017	8893.716	9010.973	8842.185	8937.603	8800.660	8941.258	8771.206
• 16	10369.190	10179.198	9950.783	9950.455	10573.680	10067.665	9979.144	9982.656
• 17	9609.738	9151.590	9125.210	9099.426	9635.448	9245.354	9086.669	9090.233
• 18	12921.281	12697.459	12736.838	12631.843	12781.325	12497.641	12636.751	12470.136
• 19	11119.323	10762.871	10322.383	10318.235	11321.046	10545.611	10190.318	10178.118
• 20	9858.545	9319.291	9239.382	9240.562	9715.703	9264.684	9252.650	9254.490

Parameter estimates for selected Q-Learning models

Table S7 presents for each rat all estimated parameters arising from models selected using BIC (see Table S5 for values of this criterion).

Table S7: Parameter estimates by rat for all selected Q-Learning models

Rat	λ_l	λ_g	α_g	α_l	γ_g	γ_l	α_g^f	α_l^f
• 1	-0.2989	1.8905	0.2167	0.0186	0.1326	-0.9778	1.0000	0.0023
• 2	-1.9655	2.5173	0.0095	0.0001	-0.2132	-0.1046	0.0095	0.0001
• 3	-0.4393	10.2292	0.0242	0.0172	0.0000	0.0000	1.0000	0.0084
• 4	-0.5987	0.2694	0.8493	0.0060	-0.1491	-0.1276	0.8493	0.0060
• 5	-1.3210	0.0853	1.0000	0.0000	-0.1456	-3.0508	1.0000	0.0000
• 6	-0.4528	2.9762	0.0884	0.0210	0.5453	-0.1583	1.0000	0.0028
• 7	-0.5457	0.8926	0.2932	0.0058	-0.4648	0.2237	1.0000	0.0001
• 8	-0.2567	0.5390	0.0474	0.0147	-0.0042	-0.2049	0.0325	0.0033
• 9	-0.3810	11.7451	0.0216	0.0088	-0.1397	0.1999	1.0000	0.0008
• 10	-0.1673	1.0972	0.6405	0.0412	0.1963	-1.6800	1.0000	0.0001
• 11	-0.1694	1.0574	0.0124	0.0509	0.0000	0.0000	0.0016	0.0047
• 12	-0.2108	0.8191	0.0047	0.0510	0.0000	0.0000	0.0002	0.0054
• 13	-0.2902	2.8084	0.0045	0.0534	0.1628	-0.5758	0.0275	0.0019
• 14	-0.0711	0.4545	0.3508	0.6407	0.2730	-3.2954	1.0000	0.8990
• 15	-0.3196	0.5887	0.0316	0.0439	0.0377	0.0938	0.0017	0.0062
• 16	-0.1964	0.0364	0.0003	0.0308	-0.1129	-2.6553	0.0003	0.0308
• 17	-4.0915	0.3238	0.7455	0.0000	-2.6254	0.1739	0.7455	0.0000
• 18	-0.3422	-0.3281	0.5824	0.0301	-0.2130	-0.3513	1.0000	0.0033
• 19	0.0548	1.7988	0.0001	0.9999	2.7104	-0.5323	0.0001	0.9999
• 20	40.1887	2.0923	0.0009	0.0000	0.0746	-3.5294	0.0009	0.0000

Comparison of mean and median parameter estimates of augmented Q-Learning models for blue and green phenotypes

Table S8 presents median and mean parameter values estimated from augmented Q-Learning models for green and blue phenotypes, as well as the relevant statistical tests. Parameters difference tests are carried out by Wilcoxon Rank Sum tests (column 4) and Mean Difference tests (column 7). Significant results are highlighted.

Table S8: Mean and median parameter estimates for blue and green phenotypes, and difference tests with significant results highlighted

Parameter	Median blue rats (•)	Median green rats (•)	Wilcoxon Rank Sum test (p-value)	Mean blue rats (•)	Mean green rats (•)	Mean Difference test (p-value)
λ_g	1.0972	0.8926	0.8490	1.9217	2.9646	0.6717
λ_l	-0.2989	-0.3196	1.0000	1.9834	-0.3345	0.3674
γ_g	0.0000	-0.0042	0.3233	0.0425	-0.1142	0.1973
γ_l	-0.5758	0.0938	0.0102	-1.2700	0.0625	0.0047
α_g	0.0884	0.0316	0.8490	0.2902	0.0813	0.2126
α_l	0.0210	0.0147	0.9241	0.0700	0.0248	0.9694
α_g^f	1.0000	0.0325	0.6331	0.6067	0.4072	0.2875
α_l^f	0.0028	0.0033	0.9241	0.0739	0.0030	0.4526

An augmented Q-Learning model with outcome range-adaptation

As a robustness check, we introduce in our augmented Q-learning models the range principle due to [33] - see also [31] - which captures the notion that subjective evaluation of rewards may take into account their range through the Min-Max normalization presented below. That is, we replace objective rewards by their subjective judgements $s(.)$ as follows:

$$s(r^k) = \frac{r^k - r^k_{\text{min}}}{r^k_{\text{max}} - r^k_{\text{min}}} \quad \text{with} \quad k \in \{g, l\}$$

Subjective judgements were then substituted in equation (1) for gains and losses and the resulting Q-learning models were estimated using the procedure described in Material and Methods, section Augmented Q-Learning Model Estimation and Simulation. More precisely, subjective judgements for gains and losses are as follows: $s(r^g) = (r^g - 1)/(80 - 1)$ for options that give convex gains and $s(r^g) = (r^g - 2)/(5 - 2)$ for options that give concave gains; similarly, $s(r^l) = (r^l + 15)/(-6 + 15)$ when losses are convex losses and $s(r^l) = (r^l + 240)/(-3 + 240)$ when losses are concave. REE are then used (as Min and Max values) to normalize all rewards and they are included in the Q-sub-values when they happen. This implies, by definition, that subjective judgements equal zero when Black Swans materialize, and equal one when Jackpots occur.

When REE are introduced through the normalization of gains and losses described above, models that consider nonzero decisions weights for REE are selected for 18 rats out of 20. Augmented Q-learning models with outcome range-adaptation proved however to improve the quality of the models for four rats only and we kept with our parsimonious models in the article. BIC and AIC values by rat are presented in the Tables S9-S10, for all four models with REE included in the Q-sub-values.

Table S9: BIC values for all augmented Q-Learning model with outcome range-adaptation, with selected value highlighted

	Model 1	Model 2	Model 3	Model 4
• 1	10360.052	9731.387	9632.944	9497.435
• 2	13380.594	13389.451	13101.613	13021.902
• 3	8867.516	8755.716	8863.804	8750.195
• 4	9668.138	9630.168	9388.315	9404.728
• 5	8974.453	8937.020	8730.574	8746.844
• 6	7323.166	7081.788	7116.992	6967.601
• 7	9921.860	9584.036	9932.662	9542.869
• 8	12817.763	12788.346	12639.371	12656.187
• 9	12750.633	12648.171	12219.681	11858.015
• 10	9565.135	9199.293	9316.995	9067.288
• 11	11299.508	11169.590	11093.347	10943.598
• 12	11255.261	11141.949	11103.222	11014.454
• 13	10563.680	10416.064	10282.824	10197.848
• 14	10072.404	9837.106	9499.525	9497.333
• 15	8921.307	8861.083	8865.852	8868.116
• 16	10637.961	10207.127	9981.039	9997.678
• 17	9803.718	9282.518	9303.068	9305.074
• 18	12827.397	12787.223	12602.462	12505.592
• 19	10673.745	10668.086	10371.852	10387.190
• 20	9796.662	9334.566	9295.024	9261.452

Table S10: AIC values for all augmented Q-Learning model with outcome range-adaptation, with selected value highlighted

	Model 1	Model 2	Model 3	Model 4
• 1	10333.417	9691.434	9592.991	9444.164
• 2	13354.126	13349.747	13061.910	12968.964
• 3	8841.442	8716.605	8824.694	8698.048
• 4	9641.717	9590.536	9348.683	9351.885
• 5	8948.411	8897.957	8691.511	8694.759
• 6	7296.884	7042.366	7077.570	6915.038
• 7	9895.410	9544.362	9892.988	9489.970
• 8	12791.534	12749.003	12600.028	12603.729
• 9	12724.012	12608.240	12179.750	11804.773
• 10	9538.739	9159.698	9277.400	9014.495
• 11	11273.547	11130.648	11054.406	10891.676
• 12	11229.062	11102.651	11063.924	10962.057
• 13	10537.316	10376.518	10243.278	10145.120
• 14	10045.630	9796.945	9459.364	9443.785
• 15	8894.609	8821.036	8825.806	8814.721
• 16	10611.321	10167.168	9941.080	9944.399
• 17	9777.224	9242.777	9263.328	9252.088
• 18	12800.603	12747.032	12562.271	12452.004
• 19	10646.914	10627.840	10331.606	10333.529
• 20	9770.154	9294.803	9255.261	9208.435

Outcomes

What are the ex-post outcomes of all choices made by the different rats over the course of the 41 final sessions? In the first left columns of Table S11, we report for each rat the number of nose pokes, the number of REE of each type the rat has experienced. We also report, for the rewards in pellets and the waiting times in seconds, the sum as well as the first four moments of the outcome per nose poke. The overall pattern that emerges from Table S11 is that a typical rats in with high Total Sensitivity group tends to have outcomes that differ from a typical rat with moderate Total Sensitivity. For gains (i.e. sugar pellets), the former's rewards have typically higher mean and variance but smaller skewness and kurtosis that the later's. This happens for instance, to an extreme degree, when we compare rat 15 (the most Anti-fragile rat) and rat 17. On the loss side, by symmetry the waiting time of the average high-sensitivity rat tends to have lower mean and variance, but larger skewness and kurtosis. These facts are consistent with the fact that a typical high-Total Sensitivity rat tend to pick a mix of exposures that is more convex both on the gain domain and on the loss domain, compared to an average low-Total Sensitivity rat.

Table S11: Ex-post total outcomes for each of the 20 rats over the final 41 sessions - summary statistics. Each row represents one rat and the color codes indicate the profile of the animal, as in Figure 2

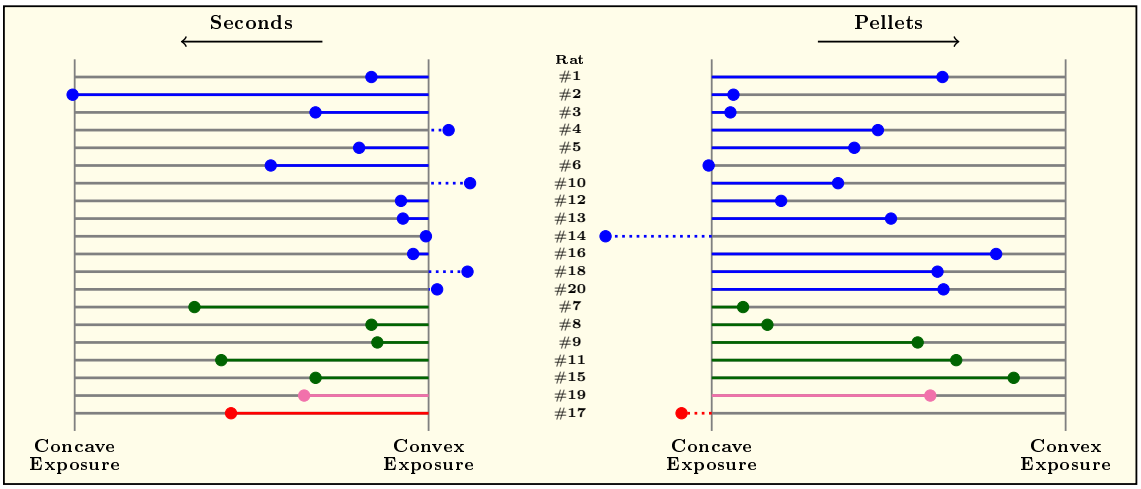
				Rewards (pellets)					Losses (seconds)				
Rat	Nosepokes	Black Swan	Jackpot	Sum	Mean	Var.	Sk.	Kur.	Sum	Mean	Var.	Sk.	Kur.
• 1	5761	3	13	8399	3.08	33.05	11.54	152.99	27408	9.04	67.54	22.12	618.13
• 2	5526	13	8	7622	2.92	23.13	12.88	203.11	28305	9.69	274.50	12.08	166.54
• 3	5006	6	4	6896	2.92	12.48	17.82	385.26	24570	9.28	147.76	15.69	294.76
• 4	5461	1	11	7833	3.04	30.05	11.97	166.29	25653	8.89	31.78	24.19	980.96
• 5	4967	3	10	7157	3.05	30.24	11.86	164.03	24036	9.16	77.56	20.79	540.67
• 6	5273	6	3	7212	2.90	9.18	20.11	507.29	25935	9.31	133.01	17.34	345.25
• 7	5499	14	6	7542	2.93	18.13	14.12	250.85	27567	9.43	312.42	10.81	139.10
• 8	5204	9	6	7258	2.95	18.42	14.37	254.07	24900	9.07	238.04	11.31	165.31
• 9	5740	10	11	8288	3.06	28.51	12.30	175.47	27417	9.05	230.45	11.89	177.40
• 10	5427	1	11	7721	3.00	30.38	11.83	163.40	25326	8.86	31.86	24.27	986.03
• 11	4867	15	12	6997	3.10	35.59	11.47	146.92	25197	9.67	353.35	10.69	129.97
• 12	5165	4	9	7092	2.95	26.97	12.41	181.89	24903	9.02	98.05	18.61	429.90
• 13	5383	2	13	7620	3.02	36.87	10.66	133.35	25761	9.02	56.28	20.88	631.19
• 14	5964	1	7	7822	2.81	20.19	13.09	219.91	28488	8.96	30.39	23.55	972.34
• 15	5851	6	18	8366	3.06	46.40	9.61	107.29	28605	9.19	125.22	17.10	349.52
• 16	5767	2	13	8187	3.04	33.07	11.69	155.14	27567	8.97	48.87	23.71	777.59
• 17	5560	14	1	7393	2.85	3.72	25.03	991.18	28095	9.47	303.48	11.09	145.03
• 18	5994	2	10	8354	2.99	24.77	13.38	205.37	28329	8.86	56.08	19.02	570.56
• 19	6049	10	19	8416	2.98	47.69	9.46	104.24	29841	9.26	200.94	13.57	218.44
• 20	5581	1	14	7927	3.04	36.55	11.20	141.46	26526	8.91	31.58	23.73	962.88

Convexity Premiums

Evidently the outcomes that we report in Table S11 depend on the convexity mix of options that each rat has chosen, as captured by our notions of Total and One-Sided Sensitivity. To go beyond Table S11 so as to capture the extent to which rats exploit convexity in their choices, it is perhaps informative to look at the outcomes for each rat in the following way. Right and left panels in Figure S1 refer to losses and rewards, respectively.

In each panel, rats are color-coded as in Figure 2, and they are represented horizontally by two segments which are normalized in the following way. Consider for example (blue) rat 1 in the first line. In the right panel about pellets, the point of the black segment most to the left represents the outcome that would have happened, had the rat chosen a concave option in the gain domain (that is, either Robust or Fragile), exclusively in all sessions. The point most to right corresponds, in contrast, to the counterfactual outcome in which all nose pokes of the rat correspond to a convex option (that is, either Vulnerable or Anti-Fragile). Superimposed on this background grey segment, is a bold-colored segment (blue because rat 1 is blue) that ends with a circle indicating what the rat has gained due to

Figure S1: Convexity Premiums for each of the 20 rats (row) - see Appendix for details. Each row represents a rat and the color code indicates the profile of the animal, as in Figure 2. Towards the right, the dot materializes how many pellets were obtained relative to the number that would have been obtained if a convex menu had been chosen at each trial. Towards the left, the coloured dot indicates the total seconds of penalty obtained relative to what would have been obtained if a concave menu had been chosen at each trial.



convexity in relative terms - what we call the *convexity premium* (see Appendix for more details). Rat 1 has a normalized convexity premium that corresponds to roughly 80% of what it would have got, in terms of sugar pellets, had it counterfactually chosen to mix only Vulnerable and Fragile, instead of Robust and Anti-fragile as it did. Similarly, one sees from the right panel of Figure S1 that the Anti-fragile rat 15 has a convexity premium that corresponds to roughly 90% of what it would have got, in terms of sugar pellets, had he chosen to mix only Vulnerable and Fragile. From panel (b) of Figure 2, we know that rat 15 has also, more rarely, picked the Robust and Fragile exposures and this is why its convexity premium in the gain domain is less than maximal. At the other extreme, the rat 6 has a convexity premium near zero. Note that two rats are outside the black segment in the gain domain, that is, have "negative" convexity premiums, and they are indicated by dashed lines. Rat 17 has picked the Anti-fragile exposure a few times but got only one extreme gain, while rat 14 is more sensitive and hence got more extreme gains but still too few of them. Therefore, both rats got less sugar pellets than what they would have gotten by mixing exposures that are concave in the gain domain (Robust and Fragile).

The left panel in Figure S1 about time-out punishments can be read in a similar way, with now the point most to the left of the black segment corresponds to the largest loss in terms of waiting times, while the end point to the right corresponds to the lowest amount of time wasted. Therefore, the bold and colored segment indicates a negative premium: for instance, rat 2 on the first line in the left panel of Figure S1 has been exposed to the largest time-out punishment: it has a negative and large convexity premium because it mixed exposures with concave losses (that, is Fragile and Vulnerable), too often. Symmetrically, some rats turn out to get convexity premiums that are above 100% because they got too few black swans when they picked exposures that are concave in the loss domain). Comparing the right and left panels in Figure S1, one infers that *rats exploit convexity better in the loss domain than in the gain domain, that is, they more often avoid Black Swans than they get Jackpots*. This is of course consistent with our earlier observation that most rats exhibit moderate to high Black Swan Avoidance.

References

[1] Adams W.K., Haar C.V., Tremblay M., Cocker P.J., Silveira M.M., Kaur S., Baunez C., Winstanley C.A. (2017). Deep-brain stimulation of the subthalamic nucleus selectively decreases risky choice in risk-preferring rats. eNeuro Aug 7;4(4). 3

[2] Ainsworth T.D., Heron S.F., Ortiz J.C., Mumby P.J., Grech A., Ogawa D., Eakin C.M., Leggat W. (2016). Climate change disables coral bleaching protection on the Great Barrier Reef. Science 352, 338-342. 27

[3] Bao C., Zhu X., Moller-Mara J., Li J., Dubroqua S., Erlich J.C. (2023). The rat frontal orienting field dynamically encodes value for economic decisions under risk. Nature Neuroscience 26(11): 1942-1952. 24, 25

[4] Bechara A., Damasio A.R., Damasio H., Anderson S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50: 7-15. 3

[5] Benartzi S., Thaler R.H. (2001). Naive diversification strategies in defined contribution saving plans. The American Economic Review, 91: 79-98. 12

[6] Breysse E., Meffre J., Pelloux Y., Winstanley C. A., Baunez C. (2021). Decreased Risk-taking and Loss chasing after Subthalamic Nucleus Lesion in the Rat. Eur. J. Neurosci. 53(7): 2362-2375. 3

[7] Charpentier C.J., De Martino B., Sim A.L., Sharot T., Roiser J. P. (2016). Emotion-induced loss aversion and striatal-amygdala coupling in low-anxious individuals. Soc Cogn Affect Neurosci. 11(4):569-79. 26

[8] Chevin L.-M., and Hoffmann A.A. (2017). Evolution of phenotypic plasticity in extreme environments. Phil. Trans. R. Soc. B 372, 20160138. 26, 27

[9] Cinotti, F., Fresno, V., Aklil, N., Coutureau, E., Girard, B., Marchand, A.R., and Khamassi, M. (2019). Dopamine blockade impairs the exploration-exploitation trade-off in rats. Nature Scientific Reports, 9: 6770. 2, 26

[10] Constantinople,C.M., Piet, A.T., and Brody, C.D. (2019). An analysis of decision under risk in rats. Current Biology, 29: 2066-2074. 2, 8, 25, 37

[11] Delgado M.R., Li J., Schiller D., Phelps E.A. (2008). The role of the striatum in aversive learning and aversive prediction errors Philos Trans R Soc Lond B Biol Sci. 63(1511):3787-800. 26

[12] Easterling D.R., Meehl G.A., Parmesan C., Changnon S.A., Karl T.R., Mearns L.O. (2001). Climate extremes: observations, modeling, and impacts. Science 289, 2068-2074. 27

[13] Ferrari-Toniolo, S., Seak, L.C.U., Schultz, W. (2022). Risky choice: Probability weighting explains independence axiom violations in monkeys. Journal of Risk and Uncertainty 65, 319â351. 25

[14] Ferrari-Toniolo S., Schultz W. (2023). Reliable population code for subjective economic value from heterogeneous neuronal signals in primate orbitofrontal cortex. Neuron. 111(22):3683-3696. 26

[15] Fishburn, P.C., and Vickson, R.G. (1978). Theoretical foundations of stochastic dominance. In "Stochastic Dominance", G. A. Whitmore and M. C. Findlay, eds. D. C. Heath and Co., Lexington, Massachusetts, 37-113. 35, 36

[16] Fletcher, R., Reeves, C. M. (1964). Function minimization by conjugate gradients. Computer Journal 7, 148-154. 42

[17] Genest W., Stauffer W.R., Schultz W. (2016). Utility functions predict variance and skewness risk preferences in monkeys. Proceeding of the National Academy of Science U S A. 113(30): 8402-7. 25

[18] Grant, P. R., Grant, B. R., Huey, R. B., Johnson, M. T. J., Knoll, A. H., Schmitt, J. (2017). Evolution caused by extreme events. Philosophical Transactions of the Royal Society B: Biological Sciences; 372 (1723): 20160146. 3

[19] Hertwig, R., and Erev, I. (2009). The description-experience gap in risky choice. Trends in Cognitive Sciences, 13: 517-523. 2

[20] Hu, Y., van Wingerden, M., Sellitto, M., Schäble, S., Kalenscher, T. (2021). Anterior cingulate cortex lesions abolish budget effects on effort-based decision-making in rat consumers. Journal of Neuroscience: 41 (20), 4448-4460. 3

[21] Kahneman D., and Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica 47, 263 25

[22] Kalenscher, T., and M. van Wingerden (2011). Why we should use animals to study economic decision making - a perspective. Frontiers in Neuroscience, 5: 1-8. 3

[23] Lak A., Hueske E., Hirokawa J., Masset P., Ott T., Urai A.E., Donner T.H., Carandini M., Tonegawa S., Uchida N., Kepecs A.(2020). Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon. Elife. 9:e49834. 3

[24] Loehman, E. (1994). Rank dependent expected utility: stochastic dominance, risk preference, and certainty equivalence. Journal of Mathematical Psychology, 38: 159-197. 37

[25] Ludvig, E.A, Madan, C.R., McMillan, N., Xu, Y., Spetch, M.L. (2018). Living near the edge: how extreme outcomes and their neighbors drive risky choice. Journal of Experimental Psychology: General, 147: 1905-1918. 26

[26] Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., and Gonzalez, C. (2015). Unpacking the exploration-exploitation tradeoff: a synthesis of human and animal literatures. Decision, 2: 191-215. 2

[27] Mokyr, J. (1992). The lever of riches: technological creativity and economic progress. OUP Catalogue, Oxford University Press, number 9780195074772. 3

[28] Needham, T. (1993). A visual explanation of Jensen's inequality. The American Mathematical Monthly, 100: 768-771. 33

[29] Nelder, J., Mead, R. (1965). A simplex method for function minimization, Computer Journal, vol. 7(4), 1965, 308-313. 42

[30] Ott T., Masset P., Gouvêa T.S., Kepecs A.(2022). Apparent sunk cost effect in rational agents. Science Advances 8(6). 3

[31] Palminteri, S., Lebreton, M. (2021). Context-dependent outcome encoding in human reinforcement learning. Current Opinion in Behavioral Sciences: 41: 144-151. 52

[32] Palminteri, S., Wyart, V., Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. Trends in Cognitive Sciences: 21(6), 425-433. 42

[33] Parducci, A. (1995). Happiness, Pleasure, and Judgment: The Contextual Theory and its Applications. Lawrence Erlbaum Associates, Inc. 52

[34] Pittaras E, Cressant A, Serreau P, Bruijel J, Dellu-Hagedorn F, Callebert J, et al. (2013) Mice gamble for food: individual differences in risky choices and prefrontal cortex serotonin. Journal of Addiction Research and Therapy S4:011. 3

[35] Prelec, D. (1998). The Probability weighting function. Econometrica, 66:497-527. 37

[36] Real, L.A (1980). Fitness, uncertainty, and the role of diversification in evolution and behavior. The American Naturalist, 115: 623-638. 12

[37] Real, L.A. (1991). Animal choice behavior and the evolution of cognitive architecture. Science, 253: 980-986. 2

[38] Redish A.D., et al. (2021). Computational validity: using computation to translate behaviours across species. Phil.Trans. R. Soc., B377: 20200525. 26

[39] Rees, M., Jessica, C., Metcalf, E., and Childs, D. Z. (2010). Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proceedings. Biological sciences, 277: 1149-1151. 12

[40] Romain A., Broihanne M-H., De Marco A., Ngoubangoye B., Call J., Rebout N. and Dufour V. (2021). Non-human primates use combined rules when deciding under ambiguity. Phil. Trans. R. Soc., B376: 20190672 25

[41] Rushworth M.F., Behrens T.E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. Nature Neuroscience. 11(4):389-97. 26

[42] Schultz W., Carelli R.M., Wightman R.M. (2015). Phasic dopamine signals: from subjective reward value to formal economic utility. Curr Opin Behav Sci. 5:147-15. 26

[43] Simon N.W., Gilbert R.J., Mayse J.D., Bizon J.L., Setlow B. (2009). Balancing risk and reward: a rat model of risky decision making. Neuropsychopharmacology 34(10):2208-17. 3

[44] Simons A. M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. Proceedings. Biological sciences, 278: 1601-1609. 12

[45] Steiner A.P. and Redish D.A. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. Nature Neuroscience. 17(7): 995-1002). 3

[46] Sutton, R.S., Barto, A.G. (2020). Reinforcement learning: An introduction (2nd ed.). The MIT Press. 6, 20

[47] Sweis B.M. et al. (2018). Sensitivity to "sunk costs" in mice, rats, and humans. Science 361: 178-181. 3

[48] Taleb, N. (2012). "Antifragile, things that gain from disorder". Random House. 3, 8

[49] Teodorescu, K., and Erev, I. (2014). On the decision to explore new alternatives: the coexistence of under- and over-exploration. Journal of Behavioral Decision Making, 27: 109-123. 2

[50] Tremblay M., Cocker P.J., Hosking J.G., Zeeb F.D., Rogers R.D., Winstanley C.A. (2014). Dissociable effects of basolateral amygdala lesions on decision making biases in rats when loss or gain is emphasized. Cogn Affect Behav Neurosci. 14(4): 1184-95. 25, 26

[51] Tremblay M., Barrus M.M., Cocker P.J., Baunez C., Winstanley C. A. (2019). Increased motor impulsivity in a rat gambling task during chronic ropinirole treatment: potentiation by win-paired audiovisual cues. Psychopharmacology (Berl). 236(6): 1901-1915. 26

[52] Vestergaard M.D., Schultz W. (2020). Retrospective Valuation of Experienced Outcome Encoded in Distinct Reward Representations in the Anterior Insula and Amygdala. Journal of Neuroscience. 40(46): 8938-8950. 26

[53] de Visser L., Homberg J.R., Mitsogiannis M., Zeeb F.D., Rivalan M, et al. (2011). Rodent versions of the iowa gambling task: opportunities and challenges for the understanding of decision-making. Frontiers in Neuroscience 5: 109. 3

[54] van Wingerden M., Marx C., Kalenscher T. (2015). Budget constraints affect male rats' choices between differently priced commodities. PLoS ONE: 10(6).

[55] Yamada H., Tymula A., Louie K., Glimcher P.W (2013). Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. Proc Natl Acad Sci U S A. 110(39): 15788-93. 3

[56] Yamamori Y., Robinson O.J., Roiser J.P. (2023). Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance. Elife. 2023 Nov 14:12:RP87720. 24

[57] Zeeb F.D., Robbins T.W., Winstanley C.A. (2009). Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. Neuropsychopharmacology. 34(10): 2329-43 25