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‖

December 3, 2025

*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 gratefully thank the Editors and Referees at *eLife* for constructive and helpful comments and suggestions. The authors would like to also thank, without implicating, Raouf Boucekkine, Andrea Brovelli, Giorgio Coricelli, John Duffy, Thibault Gajdos, Brian Hill, Tobias Kalenscher, Fabio Maccheroni, Dean Mobbs, 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 European Winter Conference on Brain Research, CERGIC, BSE Summer Forum, Institut Jean Nicod, Bocconi University, CREST, GREDEG, UC Irvine, MUSEES conference in Lyon, Comparative Psychology Research Colloquium in Düsseldorf, LNC2, IAST.

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

‡Ecole Centrale Marseille. Email: louis.matis.willem@gmail.com.

§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.

‖Corresponding author. Institute of Neurosciences Timone, UMR7289 CNRS & Aix-Marseille Université and Aix-Marseille University, CNRS, AMSE. Email: patrick.pintus@univ-amu.fr.

Abstract: Most behavioral studies in animals investigate risk using outcome probabilities larger than 10%. However, real-world Decision-Making often requires evaluating events that are both extremely unlikely and highly consequential. To address this gap, we developed an experimental and computational framework to quantify how rats detect and adapt to **rare (probability $< 1\%$) and extreme (deviation from mean > 10 standard deviations) outcomes (REE)**. Using a four-armed bandit task, animals chose between options associated with probabilistic rewards (sugar pellets) or punishments (time-out delays). Depending on the animal's choice, REE can occur or not, allowing us to probe how rats integrate information across both common and fat-tailed event distributions. Across subjects, behavior showed **restricted diversification** (typically two out of four options) and clear **sensitivity to REE**, expressed as a systematic avoidance of rare and extreme punishments ("Black Swans") combined with partial exposure to rare and extreme gains ("Jackpots"). The dominant behavioral phenotype displayed a near-complete suppression of exposure to Black Swans, whereas exposure to Jackpots remained only moderate. This asymmetric sensitivity came at a cost: these animals accepted smaller frequent gains and larger frequent losses to avoid catastrophic outcomes. To account for these behaviors, we implemented an **augmented reinforcement learning model** in which REE are weighted separately from frequent events. Fitting this model to individual behavioral data captured subjects' decision patterns far better than standard Q-learning, which systematically failed to reproduce the observed asymmetry. The best-fitting model suggests that the rat brain segregates information from the central outcome distribution and the distribution tails, assigning distinct value weights to REE during action selection. This work provides the first evidence that rodents spontaneously adapt their learning strategies to avoid extreme punishments while partially maintaining exposure to rare and extreme gains, suggesting that frequent outcomes are treated separately from rare and extreme ones through different weights in Decision-Making.

Introduction

Exploration and exploitation are fundamental components of adaptive behavior under uncertainty in both humans and non-human animals (for review see [28]). However, relatively little is known about how these behavioral processes are modulated by outcomes that are both extremely rare and highly consequential - hereafter referred to as *Rare and*

Extreme Events (REE). Although risk processing has been extensively studied across psychology, behavioral economics, and neuroscience, most laboratory decision-making tasks rely on outcome distributions in which event probabilities exceed 10% (e.g. [9], [10], [41], [53]; see also [20], [27], [32] for notable exceptions). This represents a major limitation considering that in natural environments, many species must cope with events that occur far less frequently yet exert disproportionately large negative or positive impacts. Such events have shaped evolutionary trajectories and can even lead to species-wide extinction [19], while in humans they have occasionally driven technological and cultural leaps [29]. Because their probabilities and magnitudes cannot be reliably estimated from experience, REE pose profound challenges for learning and decision-making mechanisms. It remains largely unknown whether, and how, animals - including humans - detect, integrate, or adapt to such events.

Here, we address this gap by studying rats, a species widely used in behavioral and systems neuroscience to investigate the neural substrates of cost-benefit learning and decision-making (e.g. [6], [1], [49], [51], [24], [33], [47], [21], [58]; see [23] for relevance to human decision processes). We developed a novel four-armed bandit task in which rats repeatedly interact with their environment and encounter REE with probabilities $< 1\%$. Critically, these REE produce outcome magnitudes - both gains and losses - that exceed 10 standard deviations relative to the average outcome. Standard rodent analogs of the Iowa Gambling Task typically use lower-magnitude rare events with probabilities $\geq 10\%$ ([4], [1], [37], [57]).

The central feature of our task is that it allows us to test whether rats adopt behavioral strategies that minimize exposure to extreme losses (Black Swans) while preserving the possibility of extreme gains (Jackpots) - analogous to "Anti-fragile" strategies described in conceptual work [48]. In contrast, strategies that eliminate Jackpots while maintaining exposure to Black Swans are here termed "Fragile". Operationally, complete exposure to Jackpots and complete avoidance of Black Swans defines the Anti-fragile option, whereas the inverse pattern defines the Fragile option. We additionally define the Robust option, as the one excluding all REE and Vulnerable option the one exposing animals to both positive and negative REE. This framework is compatible with evolutionary hypotheses proposing that organisms may be biased toward choice policies that enhance long-term viability under uncertainty.

Our design separates outcomes into three probabilistic domains: (i) "normal events" (NE, $\approx 90\%$ probability), (ii) "rare events" (RE, $\approx 10\%$), and (iii) REE ($< 1\%$). Re-

ward magnitudes scale nonlinearly with decreasing probability such that convex options produce increasingly larger gains as probability decreases (i.e., exposure to Jackpots) while concave options limit the possible gain within a short range whatever the probability (i.e. avoidance of Jackpots). In contrast, in the negative domain, concave options produce increasingly larger losses (exposure to Black Swans), while convex ones limit the possible losses within a short range whatever the probability. Convexity here refers not to utility curvature but to the accelerating relationship between magnitude and rarity.

Based on stochastic dominance arguments, value-maximizing agents with non-decreasing value functions should prefer concave options in the NE domain (first-order dominance), and continue to do so in the RE domain if they also display risk aversion (second-order dominance). However, in the full environment including the REE, second-order dominance reverses, and convex options become optimal because avoiding Black Swans is increasingly costly when REE fail to materialize. Thus, Anti-fragile strategies are only advantageous if REE are internally represented.

We incorporate both gains (sugar pellets) and losses (time-out punishment) to capture sensitivity across valence domains. We derive two key behavioral metrics: (1) *Total Sensitivity to REE*, defined as the combined frequency of convex choices in gain and loss domains, indexing whether rats incorporate REE into their strategy; and (2) *One-Sided Sensitivity to REE*, defined as the difference between convex gain choices and convex loss choices, quantifying in particular asymmetric Jackpot-seeking versus Black Swan avoidance.

Using 20 rats (about 6000 trials per subject over 41 sessions), we report two major behavioral findings. First, rats diversify their behavior across options but typically concentrate on roughly two of the four available strategies. Nineteen of 20 animals show moderate-to-high Total Sensitivity, indicating that most animals systematically combine some exposure to extreme rewards with partial avoidance of extreme losses. Second, 13 of 20 rats demonstrate near-complete avoidance of Black Swans while only partially seeking Jackpots, producing negative One-Sided Sensitivity and consistent choice mixtures of Anti-fragile and Robust options. This suggests differential weighting of rare positive versus rare negative outcomes.

Consistent with this interpretation, both Total Sensitivity and Black Swan avoidance increase after rats directly experience a rare extreme loss. To interpret computationally these results, we evaluated a family of augmented Q-learning models that incorporate explicit sensitivity parameters for Jackpots and Black Swans. Model comparison using information criteria identifies an architecture in which NE+RE and REE are weighted separately during

action selection, suggesting that rats assign distinct cognitive or neural salience to REE - especially negative ones - when making decisions under uncertainty.

Results

Measures of sensitivity to Rare and Extreme Events in a four-armed bandit task

Rats were tested in an operant chamber equipped with four nose-poke ports, each linked to a specific distribution of reward (sucrose pellets) and punishment (time-out delays) drawn from a fixed schedule across sessions (Fig. 1a). After training, animals completed 41 experimental sessions (20 min/session; $\approx$ 120 trials/session). As shown in Figure 1b, positive outcomes consisted of sucrose pellets, whereas negative outcomes corresponded to time-out periods during which ports were inactive, preventing pellet acquisition.

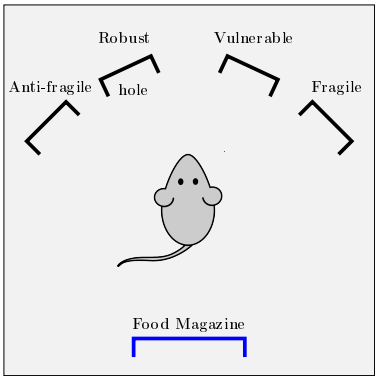
Each choice option was defined by a characteristic mapping between outcome magnitude and event probability. Outcome magnitudes could follow either convex (accelerating; green) or concave (decelerating; red) functions as event probability decreased (Fig. 1b-c). Critically, all options included a spectrum of outcome classes: highly probable Normal Events (NE; $\approx$ 90%), Rare Events (RE; $\approx$ 10%), and Rare and Extreme Events (REE; $\approx$ 1%), which corresponded to exceptionally large rewards ("Jackpots": 80 pellets) or exceptionally large punishments ("Black Swans": 240-s delay).

Combining convex or concave gain and loss functions generated four distinct options (Anti-fragile, Robust, Vulnerable, Fragile; Fig. 1c). The Anti-fragile option paired convex gains with convex losses, thereby allowing exposure to Jackpots while avoiding exposure to Black Swans. The Fragile option combined concave gains and concave losses, yielding small plateauing rewards but exposing animals to extreme losses (Black Swans). Robust and Vulnerable options showed complementary patterns: Robust avoided Black Swans but also prevented access to Jackpots, whereas Vulnerable exposed animals to both extreme outcomes.

Because the behavioral structure explicitly manipulated exposure to REE, we quantified how frequently each rat selected options containing convex (versus concave) outcome functions. Two measures were derived (Fig. 1d). Total Sensitivity to REE reflects the overall tendency to select options that allow contact with extreme outcomes (i.e., convex functions in either domain). One-sided Sensitivity quantifies asymmetry between the gain and loss

Figure 1: **Experimental Design.** (a) Schematic of the operant conditioning chamber. Rats performed a self-paced four-option choice task using nose-poke apertures, each corresponding to a distinct option leading to distribution of reward (sucrose pellets) or punishment (time-out delay) at various size and probability. (b) Outcome structure for each option. Rewards and losses follow either convex (green) or concave (red) probability-magnitude functions. Outcomes are grouped into three probability domains: high-probability Normal Events (NE; $20-25\%$, blue), Rare Events (RE; 5%, yellow), and Rare and Extreme Events (REE; $< 1\%$, pink). (c) Functional characterization of the four choice options based on exposure to rare and extreme gains ("Jackpots") and rare and extreme losses ("Black Swans"). 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. For convenience, gains appear in the upper-right quadrant while losses appear in the lower-left one. Convex curve are in green while concave are red. Robust: convex losses and concave gains (avoids both REE). Anti-fragile: convex losses and convex gains (avoids Black Swans while retaining access to Jackpots). Fragile: concave losses and concave gains (exposure to Black Swans, avoidance of Jackpots). Vulnerable: concave losses and convex gains (exposed to both Black Swans and Jackpots). (d) Behavioral sensitivity metrics. Total Sensitivity (y-axis) quantifies overall incorporation of REE by summing convex choices in gain and loss domains. One-sided Sensitivity (x-axis) measures asymmetry between REE seeking and REE avoidance. Example profiles illustrate: an Anti-fragile-like pattern (green), a Fragile-like pattern (red), and a mixed Anti-fragile/Robust profile with strong Black Swan avoidance and partial Jackpot seeking (blue). See Material and Methods for computational definitions.

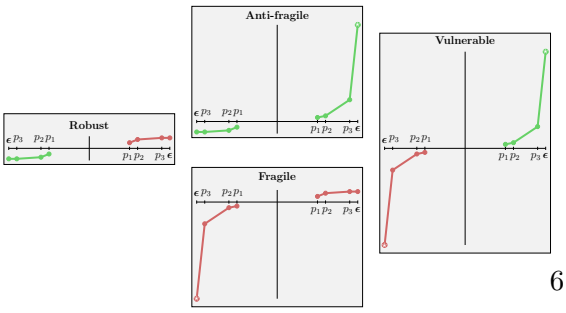
(a)



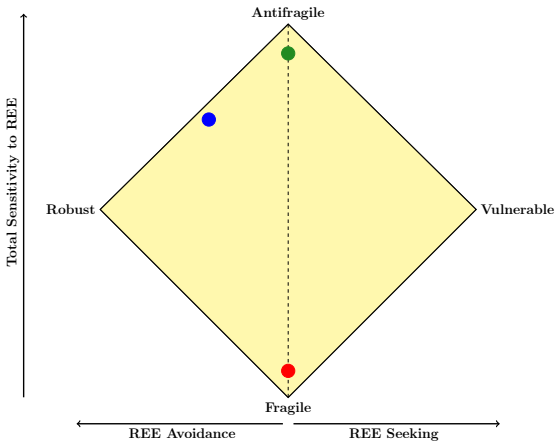
(b)

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)



domains - specifically, whether an animal preferentially seeks extreme rewards (Jackpot Seeking) or preferentially avoids extreme punishments (Black Swan Avoidance).

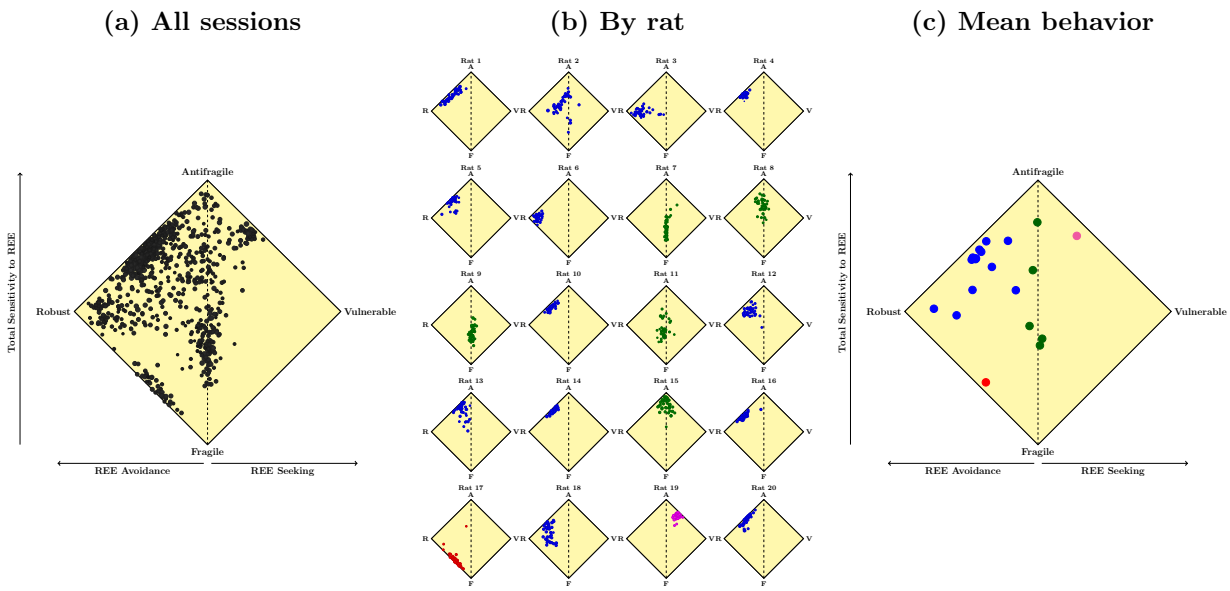
These two dimensions generate a behavioral space in which animals may cluster near extreme phenotypes (e.g., exclusively Anti-fragile or exclusively Fragile) or diversify among several strategies. Points lying between edges reflect mixed strategies. For example, a profile near the Anti-fragile - Fragile axis indicates strong Black Swan Avoidance combined with frequent-but not exclusive - Jackpot seeking.

Importantly, although Normal and Rare Events dominate the experience of the task, the inclusion of REE changes the adaptive landscape: extreme outcomes, despite their low probability, produce large shifts in reinforcement history. Thus, rats that behave as though REE "matter" (high Total Sensitivity) preferentially select options allowing contact with large rewards while minimizing catastrophic time-outs, whereas rats that behave as if REE are irrelevant (low Total Sensitivity) preferentially sample concave options that yield stable but limited outcomes, according to the stochastic dominance of order 1. Together, this framework allows quantification of how individual animals integrate extremely infrequent but high-impact outcomes into their decision strategy, enabling downstream analyses linking these behavioral phenotypes to neural processes.

Main phenotype: strong avoidance of extreme punishments with partial sampling of extreme rewards

We analyzed the behavioral data collected across 41 daily sessions from 20 rats (see Material and Methods). To highlight the dominant features of the behavioral strategies adopted by the animals, Figure 2 plots each rat's choices within the two-dimensional space defined by Total Sensitivity to REE and One-sided Sensitivity (relative weighting of extreme losses vs extreme rewards). This representation corresponds to the rotated square introduced in Figure 1d, where each vertex corresponds to exclusive selection of one of the four outcome distributions. A first inspection of Figure 2 shows that no animal operates at a single vertex - i.e., no rat commits exclusively to one contingency. Instead, all animals sample from multiple options, producing distributed behavioral profiles rather than categorical strategies. This diversification is qualitatively consistent with what is often observed in naturalistic foraging scenarios in which animals hedge against uncertainty or extreme events ([5], [40], [43], [48]).

Figure 2: Total sensitivity to Rare and Extreme Events (REE) (y-axis) plotted against one-sided sensitivity to REE (x-axis), where rightward movement indicates REE seeking and leftward movement indicates REE avoidance. For a detailed description of the modeling and statistical analysis methods used, please refer to the Materials and Methods section.
(a) the 820 black dots represent data from 41 sessions conducted with a total of 20 rats.
(b) individual rat profiles, with each square corresponding to a single rat. The colored dots within each square denote the performance across the 41 sessions for each rat, categorized into two primary sensitivity profiles: blue dots indicate rats that exhibit strong sensitivity to REE and are classified as "Black Swan avoiders", while green dots indicate rats that are responsive to REE but maintain a neutral stance, situated near the midline of the graph. Additionally, two outlier data points are highlighted in red and pink.
(c) color-coded dots, each representing the average sensitivity measures across all 41 sessions per rat (n=20 rats).



Population-level patterns

Panel 2a shows the full dataset ($\approx$ 800 session-points). Most sessions lie on the left half of the behavioral space, corresponding to strategies that rather avoid REE. Within this domain, most sessions fall in the upper quadrant, indicating moderate to high Total Sensitivity - that is, rats systematically adjusted their choices in ways that reflect the presence of rare and extreme outcomes by minimizing exposure to extreme losses (Black Swans). Training history contributing to these patterns is detailed in Material and Methods.

Individual behavioral profiles

Panel 2b displays each individual rat's pattern across the 41 sessions. Each square corresponds to one rat, and the four vertices correspond to the four possible outcome distributions (A: Antifragile; R: Robust; V: Vulnerable; F: Fragile). With the exception of one animal (rat 17, red), all rats show moderate to high Total Sensitivity. Within this dominant group, two sub-groups emerge (blue and green). A second outlier, rat 19 (purple), is the only animal expressing both very high Total Sensitivity and strong Jackpot Seeking, achieved by alternating mainly between Antifragile and, less frequently, Vulnerable options. The remaining 18 animals fall into two stable phenotypes:
(1) Blue group (13 rats): These animals primarily oscillate between the Robust and Antifragile options. This phenotype combines (i) moderate to high Total Sensitivity with (ii) pronounced avoidance of Black Swans. In other words, these rats preferentially avoid extreme punishments more than they seek extreme rewards.
(2) Green group (5 rats): These rats primarily alternate between the Fragile and Antifragile options. They show moderate to high Total Sensitivity but display no directional bias in One-sided Sensitivity, indicating neither consistent avoidance of extreme losses nor active seeking of extreme rewards. Their choices are symmetric with respect to REE in the gain and loss domains.

Average phenotypes

Figure 2c summarizes each rat's average metrics across sessions. All but four rats exhibit an average Total Sensitivity above half of the maximum value; three are slightly below one, and one rat (rat 17) sits around 0.5, making it the closest to a Fragile-like pattern, though still deviating from true Fragile behavior. In total, 19 of 20 rats show clear sensitivity to the presence of REE.

Panel 2c also reveals a pronounced dominance of Black Swan Avoidance (negative One-sided Sensitivity) in the majority of animals. Rats 17 and 19 are opposite outliers: rat 17 has low Total Sensitivity, whereas rat 19 shows the highest Total Sensitivity coupled with strong Jackpot Seeking.

Within the 18-rat majority, the two main phenotypes described above are clearly distinguished: the green subgroup shows no One-sided bias, while the blue subgroup shows pronounced Black Swan Avoidance. Quantitative values are provided in Supplementary

Material: Behavioral Measures per Rat.

Structure in the relationship between Total and One-sided Sensitivity

A notable feature in panel 2c is a positive relationship between Total Sensitivity and One-sided Sensitivity among the 16 rats with the highest Total Sensitivity. Within this high-sensitivity group, animals that more strongly incorporate REE into their choices tend to show reduced Black Swan Avoidance (i.e., their One-sided Sensitivity becomes less negative). This shift reflects a progressive tilt toward convexity in the gain domain relative to the loss domain. Rat 19 represents an extreme case, approaching the maximal combination of Total Sensitivity and Jackpot Seeking.

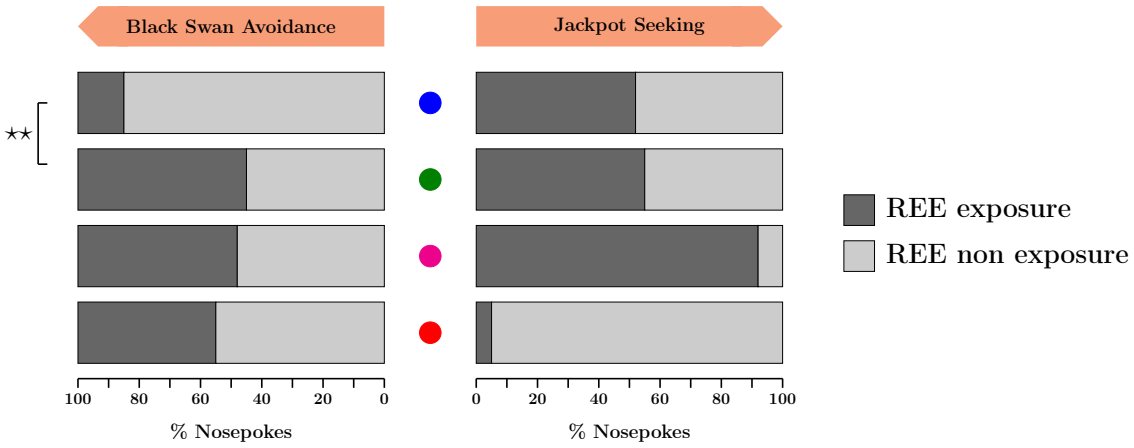
This relationship is confirmed statistically: in the full dataset of 20 rats, the correlation between average Total and One-sided Sensitivity is near zero. However, restricting the analysis to the 16 high-sensitivity rats yields a correlation of ≈ 0.57. Over learning, this relationship strengthens: splitting the 41 sessions into four blocks of 10, the correlation increases from ≈ 0.17 initially to ≈ 0.56, ≈ 0.71, and ≈ 0.64 in successive epochs. Thus, the structure linking Total and One-sided Sensitivity emerges early and stabilizes over time.

Behavioral measures of reward seeking and punishment avoidance

Figure 3 reports more intuitive behavioral metrics: the proportion of choices that (i) avoid Black Swans and (ii) seek Jackpots (detailed computation in Materials and Methods: Modeling and Statistical Analysis). Median values show a sharp dissociation between the two phenotypes. Approximately half of the blue rats avoid extreme punishments in $> 90\%$ of trials, whereas only $\approx 40\%$ of green rats-and none of the outliers-achieve comparable avoidance.

By contrast, Jackpot Seeking does not differ between blue and green rats. Wilcoxon paired tests confirm this: blue rats differ significantly between punishment avoidance and reward seeking ($p = 0.0002$), whereas green rats do not ($p = 0.6250$). The two outliers diverge in opposite directions: rat 17 seeks almost maximal Jackpot exposure, while rat 19 nearly eliminates it. Exact values are given in the Supplementary Material.

Figure 3: **Behavioral Differences Between Phenotypes in Avoidance and Seeking Behaviors.** This figure illustrates the median behavioral measures of Black Swan Avoidance and Jackpot Seeking across the two phenotypic groups, blue and green. The phenotypes are defined by distinct responses to Rare and Extreme Events (REE), with the blue phenotype exhibiting higher sensitivity to REEs and a tendency towards avoiding Black Swans, while the green phenotype shows less bias and remains closer to neutral behavior. Left Panel: This panel depicts the percentage of nosepokes where rats were exposed (black) or not exposed (light gray) to Black Swans. Increased Black Swan Avoidance is indicated by a higher proportion of non-exposure, suggesting a behavioral preference to avoid aversive outcomes among the blue phenotype compared to green. Right Panel: Conversely, this panel shows the percentage of nosepokes where rats were exposed (black) or not exposed (light gray) to Jackpots. Enhanced Jackpot Seeking is indicated by a higher proportion of exposure, suggesting greater motivation towards reward acquisition in the blue phenotype compared to green. Statistical comparisons between the phenotypes were conducted using Wilcoxon Rank Sum tests. Significant differences are denoted with $\star$ ($p < 0.05$) and $\star\star$ ($p < 0.01$).

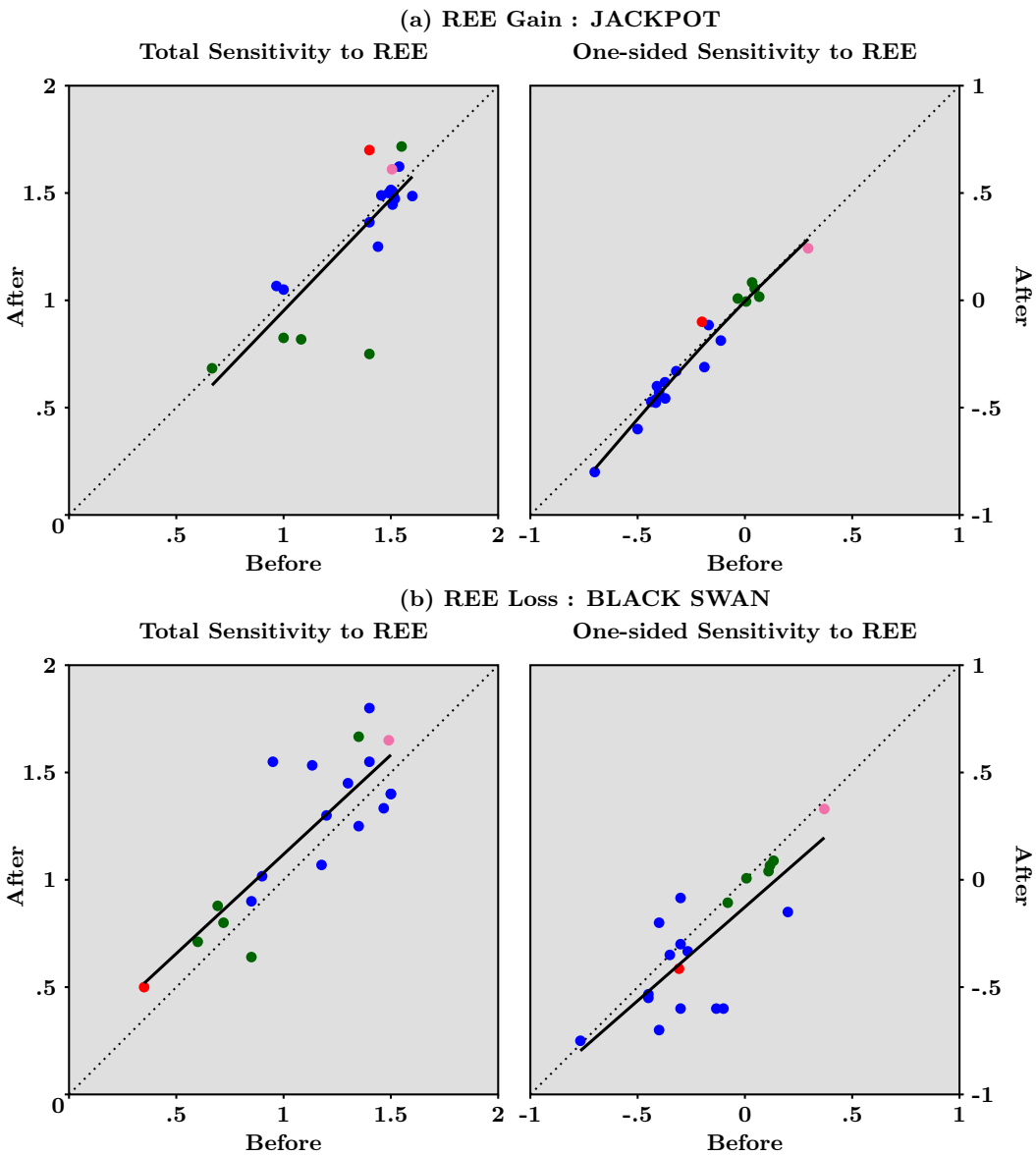


Black Swan outcomes strengthen sensitivity to rare and extreme events, whereas Jackpots do not

After establishing that most animals (19/20) display moderate-to-high Total Sensitivity to REE and robust avoidance of extreme punishments (Black Swans), we next asked how the actual experience of an REE shapes subsequent choice behavior within the same session. Specifically, we compared each rat's behavioral sensitivity in the 10 nose-pokes preceding an REE with the 10 nose-pokes following it.

Figure 4 summarizes these within-session dynamics. For each rat (color-coded as in Figure 2), panel 4a shows the effects of experiencing a Jackpot, whereas panel 4b shows the effects of encountering a Black Swan. Total Sensitivity is plotted on the left column, and

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.



One-Sided Sensitivity (relative weighting of extreme losses vs gains) on the right. The dotted diagonal indicates no change.

A clear asymmetry emerges: Jackpots produce minimal change in either metric, with most points lying close to the diagonal. In contrast, Black Swan outcomes shift behavior systematically. Most animals show an increase in Total Sensitivity - i.e., they become more responsive to the structure of rare events-along with a decrease in One-Sided Sensitivity, reflecting a further shift toward avoidance of extreme losses. In both measures, points cluster above (Total Sensitivity) or below (One-Sided Sensitivity) the 45° line, indicating a consistent strengthening of REE-related behavioral weighting after punishment but not after reward.

Thus, rare and severe losses - but not rare gains-acutely potentiate the animals' sensitivity to REE. This pattern aligns with our overall finding that rats prioritize avoidance of extreme negative outcomes and that negative REE have disproportionate influence on choice updating. Bootstrapped statistics and additional tests reported in Material and Methods (see Table 3) confirm the significance of these effects.

In summary, within-session analyses reveal a strong valence asymmetry: Black Swan events transiently enhance both global sensitivity to REE and the bias to avoid rare losses, whereas Jackpots do not measurably alter behavioral sensitivity. This suggests that rare aversive outcomes exert a privileged influence on decision-making processes in this task, reinforcing the dominant phenotype of heightened vigilance toward extreme negative events.

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:

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

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 should be thought of as averages - though not arithmetic but of an exponential moving type in view of equations (1) - see page 32 in [50]. 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 $\mathbb{1}_{JP}(o)$ equals one if the option exposes to Jackpots (namely when the chosen option is either A or V) and zero otherwise. Similarly, $\mathbb{1}_{BS}(o)$ equals one if the option exposes

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	• $\{4; 16; 20\}$

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 our model-fitting procedure (described in detail below), the key parameters of interest are the decision-weight terms for rare and extreme gains (γ_g) and losses (γ_l). In the baseline Q-learning model, both terms are fixed at zero, meaning that REE do not receive any special weighting beyond their contribution to standard value updates. We therefore compared two families of candidate models to this benchmark: one in which REE contribute directly to option sub-values, and another in which REE influence choice only through dedicated decision-weight parameters (γ_g,γ_l).

The latter class of models is especially informative. If selected, it suggests that animals treat REE as qualitatively distinct inputs to the decision process-processed separately from average reward estimates-rather than integrating them into conventional value updating. Overall, the augmented Q-learning framework includes up to eight free parameters: learning rates for gains and losses, forgetting rates for gains and losses, decision weights for gain- and loss-related sub-values, and the two REE-specific decision weights γ_g and γ_l.

Model comparison using information criteria (summarized by phenotype in Table 1) shows that 17 of 20 animals incorporate REE using dedicated decision weights, whereas only three animals are best fit by models in which REE carry no special computational status. Most rats (including 8 blue, 4 green, and the red and pink animals) rely specifically on γ_g and/or γ_l to shape choices, with or without distinct forgetting parameters. A minority of blue rats incorporate REE both through decision-weight parameters and through REE contributions to sub-values. Learning and forgetting rates did not systematically distinguish between model classes.

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

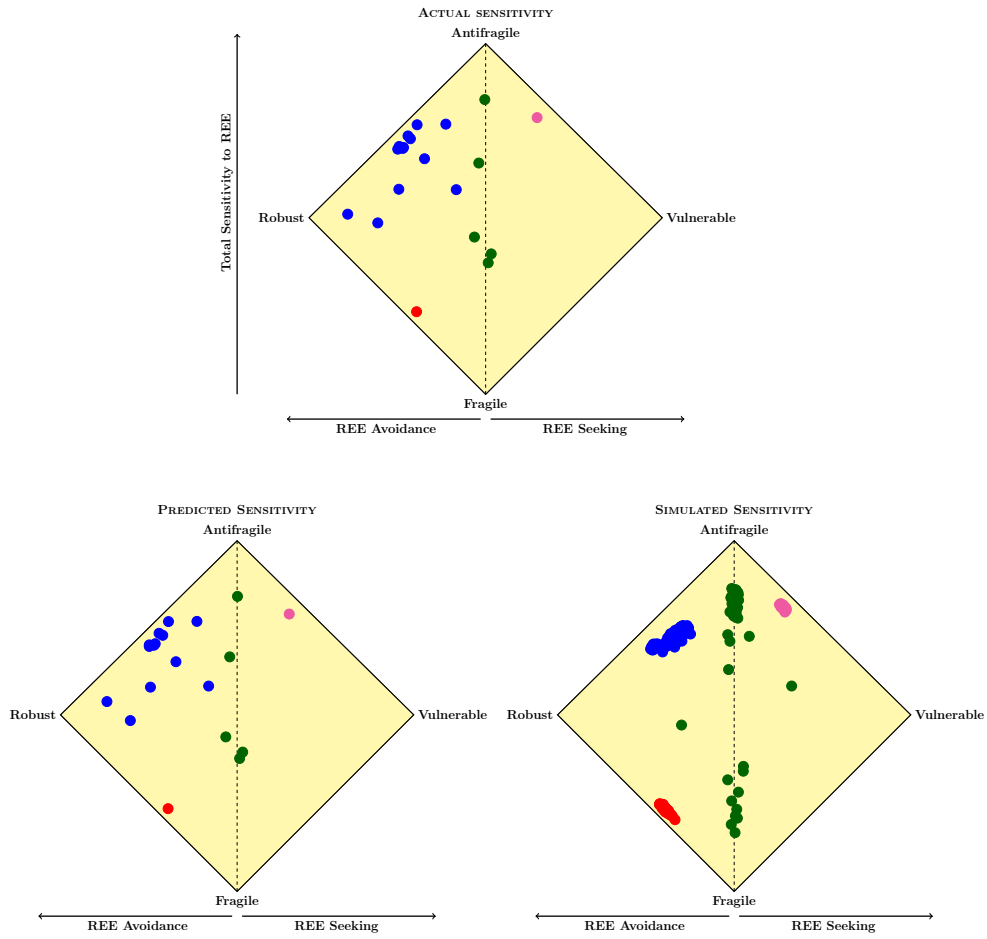
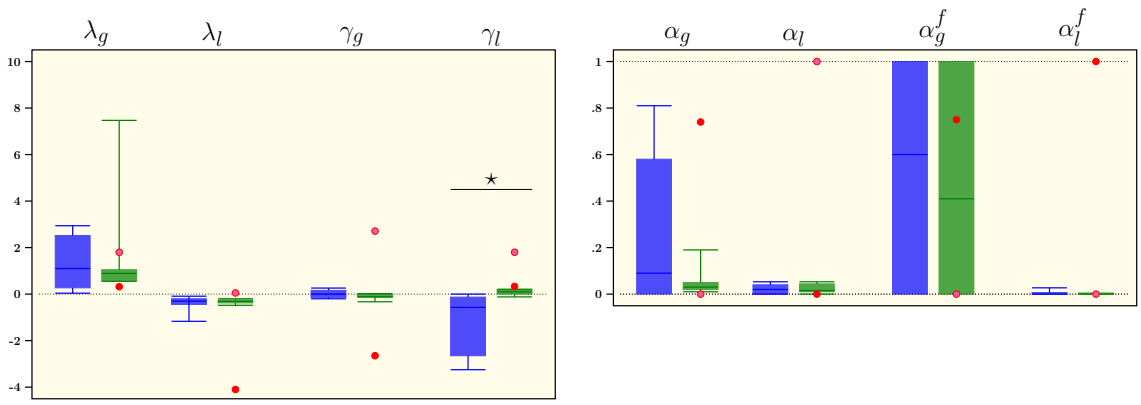


Figure 5 presents the predicted sensitivities (bottom left) and simulated sensitivities (bottom right), compared with empirically observed sensitivities (top; identical to Figure 2, right panel). Predicted sensitivities - computed using each rat's estimated parameters and the actual reward sequence - closely reproduce empirical sensitivity patterns. Simulated sensitivities were obtained by running each selected model through artificial 41-session datasets, using median parameters for the blue and green phenotypes (and individual parameters for the red and pink rats). These simulations show that the selected models generate stable behavioral phenotypes for the blue, red, and pink animals. Green rats, however, tend to segregate into two stable subgroups along the vertical (One-Sided Sensitivity) axis.

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



Estimated parameters for each rat are shown in Figure 6. A prominent distinction between phenotypes emerges when focusing on the REE-specific decision weights: blue rats show a strong selective sensitivity to Black Swan, expressed as significantly negative γ_l values ($p = 0.0102$). This indicates a robust computational penalty assigned to options that expose the animal to Black Swans (options F and V), consistent with their observed behavioral avoidance of these outcomes. In contrast, green rats do not show this selective aversive weighting and typically combine options F and A, suggesting a less selective treatment of REE. The before/after analysis illustrated in Figure 4 shows that the behavioral response to Black Swans is locally small in terms of both Total and One-sided sensitivities. This suggests that such effects are likely to be too subtle to be captured by this class of models for most rats.

Discussion

Asymmetric behavioral responses to rare extreme losses and rare extreme gains

We showed, first, that all rats in our sample exhibit some degree of Total Sensitivity to REE, with most animals displaying medium to high sensitivity. Second, we identified a dominant behavioral phenotype characterized by near-complete Black Swan Avoidance combined with partial Jackpot Seeking. We interpret these findings as reflecting a fundamental asymmetry between negative and positive REE, with a few caveats. Though we cannot fully exclude the possibility that satiety contributed to the lack of complete Jackpot Seeking, we did not systematically observe animals that stopped seeking the Jackpot. Because we restricted the rare outcome to occur on either the 10th or the 60th activation in a session (Table S1 in Supplementary Material), the question of whether the animals learn this association may arise. If the animals had learnt the 10th and 60th activation, they would exhibit a choice strategy that would tend to be more optimized than what is observed. For example, the options offering the possibility to obtain the Jackpot are not optimal in terms of gains for the frequent events, therefore the animals should tend to select these options only around the 10th and 60th choice. Most of their other choices should favor the options delivering the larger gains in the frequent domain. This is not what is observed.

In our task, avoiding Black Swans requires accepting longer frequent delays, and seeking Jackpots requires foregoing larger frequent gains. However, rats can guarantee the avoidance of Black Swans by selecting convex loss schedules, making these options relatively attractive. In contrast, no sequence of choices can guarantee the occurrence of a Jackpot-its realization depends on rare stochastic draws. For illustration, if 150 choices are made in a session (within the observed range) and the ex-ante REE probability is 1%, a Poisson approximation gives a probability of $\approx 78\%$ for observing exactly one REE-a non-negligible chance for zero occurrence. Thus, our design captures a general asymmetry likely to characterize positive versus negative REE more broadly: while extreme losses can in principle be fully avoided, extreme gains cannot be assured, even under persistent optimal seeking.

This asymmetry has important consequences for choice behavior. Rats experience a continuous stream of evidence-derived from the frequent-outcome domain-that points toward concave options as locally optimal. In particular, the Fragile option yields higher frequent

gains and smaller frequent losses, and stochastically dominates the other options when REE do not occur. Although Fragile is dominated by Anti-fragile when both the Black Swan and Jackpot occur, complete Black Swan Avoidance provides near-certain protection against extreme losses. This makes strategies combining Anti-fragile and Robust options attractive, despite violating Stochastic Dominance in the frequent domain: full-domain Stochastic Dominance is effectively secured.

In contrast, complete Jackpot Seeking is less attractive because it cannot ensure the occurrence of a Jackpot; the cost of receiving smaller gains on most trials becomes unjustified unless the compensatory extreme reward actually occurs. Partial Jackpot Seeking therefore emerges as a more balanced strategy: it reduces violations of Stochastic Dominance in the frequent domain while still allowing some exposure to large rare gains. Put simply, bearing frequent relatively large losses is acceptable only if the avoided Black Swan is certain, and bearing frequent small gains is acceptable only if the compensating Jackpot is likely to appear-which it is not. Our results indicate that the "blue" rats have learned this property, combining Robust and Anti-fragile options to achieve near-complete Black Swan Avoidance with partial Jackpot Seeking. In contrast, "green" rats continue to mix Anti-fragile and Fragile options, leaving them exposed to Black Swans.

Interestingly, since rats do not stick to one choice but show some flexibility in their strategy, this rules out the influence of a spatial bias as the main driver in their response. Interestingly indeed, the design of the task was biased in two ways: the favorite hole during training was allocated to the antifragile option as a proof of concept and the REE occurred at the 10th and 60th choice if the appropriate option was chosen. Since most rats diversified their choices, the biased location does not seem to have prevented rats to learn all options. One might wonder whether or not they may have understood and learnt the 10th and 60th choice REE delivery. If this was the case, a trend towards optimal choice such as options delivering the largest rewards and the lowest losses in the most frequent domain with a shift towards options with positive REE and avoidance of negative REE around the 10th and 60th trial should have been observed. This was never the case.

In theory, because frequent outcomes occur often, animals should be able to detect Stochastic Dominance in the frequent domain relatively quickly. However, prior work indicates that animals require many sessions to learn Stochastic Dominance (e.g. [59], [3]). Our findings suggest a possible explanation: substantial experience may be needed to infer that REE have not occurred and can be excluded from the estimated distribution.

At this stage, it is natural to ask whether our findings reflect loss aversion or ambiguity aversion. Although most studies do not focus on REE, tasks involving both positive and negative outcomes generally report loss aversion (e.g., [18], [14], [61]), consistent with Prospect Theory ([22]). Models based on this framework, however, tend to fit rat behavior more effectively in tasks involving only rewards, where losses correspond to the absence of a positive outcome rather than punishment ([10],[3]). Emotional and anxiety-related mechanisms may also contribute ([60]), potentially implicating the basolateral amygdala, which has been linked to differential evaluation of negative outcomes in rats ([54]). Ambiguity aversion in non-human animals remains relatively unexplored (but see [44]).

Our results tentatively suggest that rats may exhibit ambiguity reduction with respect to REE: combining near-complete Black Swan Avoidance with only partial Jackpot Seeking collapses the sampling probability of Black Swans toward zero, while keeping the sampling probability of Jackpots bounded between zero and ε. This behavioral asymmetry provides further evidence that extreme gains and losses receive distinct treatment during decision making. Classical Q-learning models, which allow differential weighting of gains and losses, fail to capture these patterns. In contrast, the augmented Q-learning models incorporating explicit sensitivity to REE accurately reproduce the observed behaviors - principally through the strong aversion to Black Swans that distinguishes the "blue" phenotype from the "green" one.

While a wavy recency effect has been documented in human experiments ([38]), the before/after analysis reported in Figure 4 suggests that there is no sizeable immediate recency effect for Jackpots in rats. Even for Black Swans, the immediate recency effect we report remains modest when using a 10-trial window, and the analysis of the choice immediately following a REE does not show evidence of immediate negative recency. Although further investigation is beyond the scope of this paper, possibly interacting recency effects of Jackpots and Black Swans seems a promising research avenue. Also note that the task design does not allow rats to completely avoid negative rare events (RE) unless they cease performing the task altogether - a pattern typically seen in paradigms involving aversive stimuli such as electric foot shocks. The fact that all 20 rats maintained stable performance across the 41 sessions therefore provides evidence against a pronounced "hot-stove effect" (see [12]).

The augmented Q-learning models suggest a specific neural pathway for REE encoding

The close fit and predictive accuracy of the augmented Q-learning model for the dominant behavioral phenotype point to a novel neurobiological hypothesis. Within this modeling framework, 11 of the 13 rats in the blue group (as defined in the Materials and Methods) appear to exclude REE from the Q-value update, yet display a specific sensitivity to Black Swans through a negative parameter γ_l. In other words, when evaluating available actions, "blue" rats do not integrate Black Swans into the running average of losses; instead, they assign them a distinct decision weight, effectively reducing the subjective value of options that expose the animal to these extreme losses.

In humans, rare and extreme stimuli are known to be disproportionately salient in perception and memory compared to moderate ones (e.g., [26]). Our behavioral and modeling results suggest that rats may exhibit a similar cognitive bias. This, in turn, raises the possibility that REE - particularly rare and extreme losses-may rely on neural encoding mechanisms that differ from those that process frequent, small rewards, which are classically attributed to dopamine neurons ([9], [46], [61], [55]), the striatum ([7], [11]), and cortical regions such as the anterior cingulate and orbitofrontal cortex ([45], [15]).

Identifying where and how REE are encoded in the brain remains an open question. If integrating REE requires representing long-term or low-frequency outcomes-beyond the short-term reward horizons typically studied (e.g., [56]) - candidate regions may include the anterior insula and the amygdala. The latter, given its central role in emotional salience, is a plausible substrate for representing highly aversive and unpredictable outcomes, which are known to bias choice asymmetrically toward negative events ([7], [54]).

From a computational standpoint, it is statistically rational to avoid folding extreme outliers into the mean of a distribution, as this can destabilize the estimate. Analogously, adaptive decision-making may benefit from treating REE separately from frequent outcomes, and from imposing different weights on Black Swans versus Jackpots. If such differential encoding confers evolutionary advantages in volatile environments (see [8]), then we might expect dedicated neural mechanisms to have evolved-possibly across multiple species.

To our knowledge, this specific hypothesis, grounded jointly in behavioral data and computational modeling, has not yet been directly tested in neuroscience. Doing so would align with the computational-validity framework advocated in [42], and represents an important direction for future research.

Relevance for humans

Whether humans would also find Black Swan Avoidance and Jackpot Seeking attractive - and how they might combine these strategies in a similar experimental context-is an obvious question raised by our results. Our experimental design can be readily adapted to human participants, allowing us to test whether people likewise tend to avoid harmful REE and pursue beneficial ones. We have started exploring modified designs for human subjects in which the rare but non-extreme outcomes are removed, in line with [27] who emphasize that combining rarity and extremity is key. Preliminary results indicate that the behavioral phenotypes observed in rats also emerge in humans under these modified conditions, suggesting that REE are the primary drivers of observed behaviors. Furthermore, such an extended task may be a promising tool to delineate decision-making profiles in professional environments (e.g. firemen) and also to investigate vulnerability to addiction in patients. This translational potential makes the present findings in rats relevant to broader issues concerning the past and future of humanity.

Within the framework of the Anthropocene, mounting evidence shows that climate change and other environmental disruptions driven by human activities generate extreme events (see e.g., [13]), including rare but potentially catastrophic ones that could threaten many species with partial or complete extinction. Although uncertainty remains about the magnitude of these risks, empirical data place humanity in a situation that raises difficult but pressing questions. Why has our species historically failed to avoid behaviors and policies that increase exposure to destructive REE? And to what extent will humans be able to cope with such events in the near future, should they occur? Put differently, are humans following trajectories that diverge from Black Swan Avoidance rather than moving toward it?

While it is intuitive to assume that avoidance of destructive REE-and the ability to cope with them when they occur - provides adaptive value, compelling evidence for such evolutionary mechanisms remains scarce (but see [8]). If correct, this idea introduces a paradox: humans disproportionately contribute to creating environments rich in harmful REE. Does our species possess a particular trait that impairs our ability to protect ourselves against them, despite the apparent evolutionary benefits of doing so? Because destructive REE are increasingly endogenously generated by human activity, one may even wonder whether a uniquely human developmental or epigenetic factor predisposes us to decisions that fail to eliminate exposure to extremely harmful outcomes. If so, such a factor could also impair

our ability to cope with REE once they become unavoidable. For instance, humans might struggle to detect predictive cues of REE - such as convexity in our experimental design (see [2]).

The results from our rat experiments fortunately do not support this pessimistic view. The substantial individual variability we observe-where some animals fail to avoid rare but extreme losses even when given the opportunity - suggests that humans might share this variability rather than being uniquely deficient. Importantly, the majority of rats consistently chose options that nearly eliminated exposure to harmful REE while still allowing partial access to beneficial REE in the form of Jackpots. This "adaptive" subgroup constitutes the largest portion of our sample.

A speculative hypothesis follows: some humans, and perhaps many, may also adopt strategies that avoid harmful - and sometimes pursue beneficial - rare extreme outcomes. More broadly, if sensitivity and behavioral responses to REE vary across individuals within a species, this raises important questions about how such heterogeneity shapes population dynamics, social structures, and, crucially, 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 Table 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 available 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

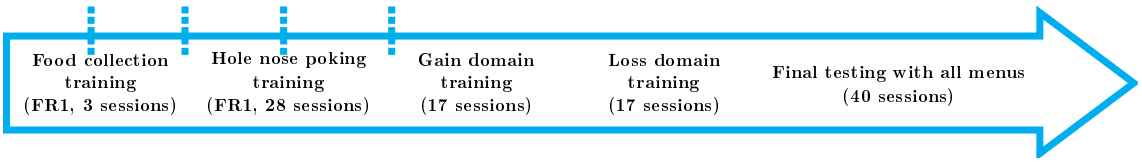
sessions, with unlimited number of pellets, during which only one hole was available (eight training sessions in total). For Anti-fragile and Vulnerable options, we used the two sequence-types where the jackpot was available at the 10th and 60th activations (Supplemental Table S1) to ensure that all 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 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 Table S2).

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

Table 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

Figure 7: Time-line of experiments



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 [30] 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.

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}

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, 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 [16]; 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 [16]). 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 [39] 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 [25] 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 [39] 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

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

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.

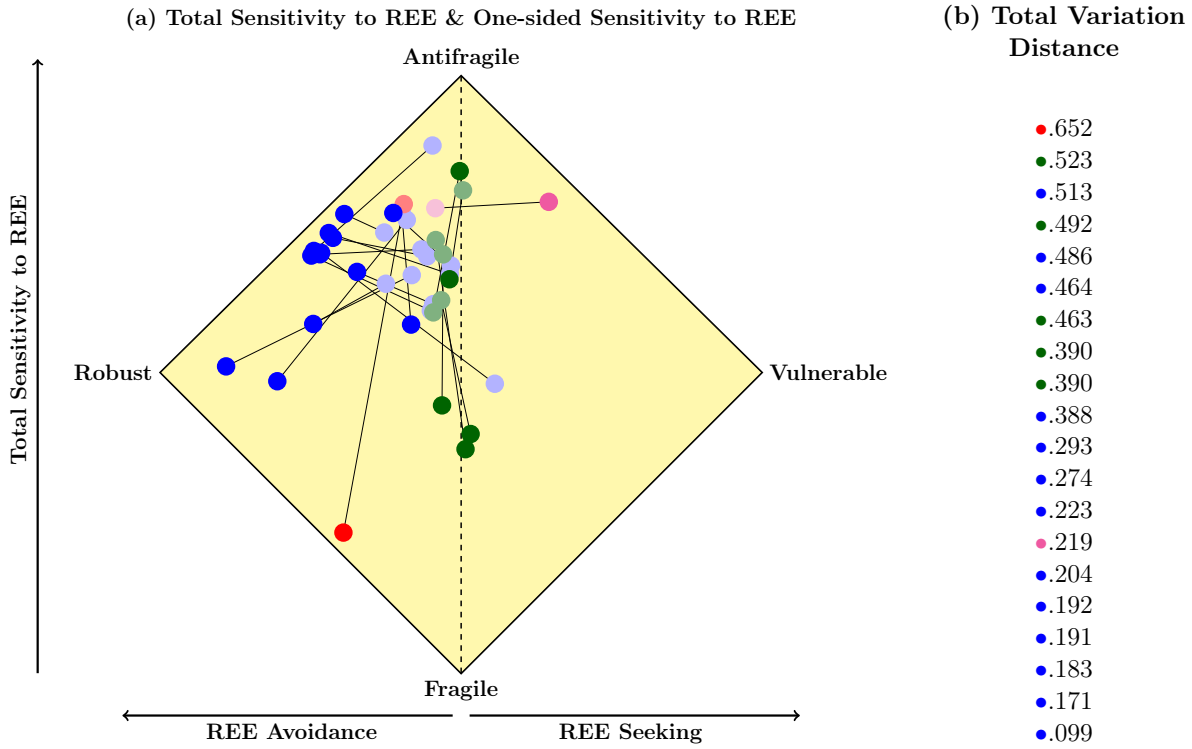
In Figure 4 we report the short term behavioral responsiveness of rat to REE. Behavioral responsiveness is measured by calculating for each rat before/after differences in TSREE

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 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. We assessed short-term behavioral responses using a paired bootstrap test (boot.paired.bca) in R (wBoot package). This non-parametric method estimates the mean difference between paired conditions by resampling with replacement and computes a bias-corrected and accelerated (BCa) confidence interval, providing robust inference without assuming normality. 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

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



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.

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 [31], 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 [17]. 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 addi-

tional 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 [35]). 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 needed to replicate the results are available via GitHub at https://github.com/PatrickPintus/Rare-and-Extreme-Events. Code is available from the authors upon request .

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,-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,-1,1,-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,-2,-2,-2,-1,-1,-1,1,-1,-1,1,1,-1,2, 1,1,2,2,-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,-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,-1,2,3,-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,-1,2,-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,-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,-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,-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,-2,1,1,1
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,-2,1,3,1,-2,-1,2,1,-2,-1,-1,-1,3,1,1,-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. The last four columns present the percentage choice of each option.

Rat	TSREE	OSREE	BSA	JPS	Anti-fragile	Robust	Vulnerable	Fragile
• 1	1.4136	-0.4583	93.59	47.76	44.76	49.69	03.93	01.61
• 2	1.2177	-0.1392	67.84	53.92	44.27	22.82	06.68	26.21
• 3	0.9749	-0.5808	77.78	19.70	09.08	70.11	09.60	11.20
• 4	1.2264	-0.4259	82.61	40.02	43.66	52.21	01.68	02.43
• 5	1.4208	-0.4990	95.99	46.09	45.36	49.11	01.43	04.07
• 6	1.0819	-0.6024	84.21	23.97	10.57	80.32	01.80	07.28
• 7	0.4632	0.0122	22.55	23.77	35.50	02.49	03.98	58.02
• 8	1.3706	-0.0669	71.87	65.18	17.39	13.46	17.34	51.79
• 9	0.6928	0.0304	33.12	36.16	32.89	05.67	09.02	52.40
• 10	1.4652	-0.4058	93.55	52.97	49.49	45.49	02.73	02.27
• 11	0.9074	0.0104	44.85	45.89	34.67	12.53	05.80	46.99
• 12	1.4365	-0.2772	85.68	57.96	47.79	37.46	02.26	12.47
• 13	1.5444	-0.1653	85.48	68.95	58.94	31.91	09.13	00.00
• 14	1.5504	-0.3937	97.20	57.83	56.61	40.59	01.28	01.50
• 15	1.6610	0.0447	80.81	85.28	75.86	08.43	08.52	07.17
• 16	1.4076	-0.5050	95.63	45.13	43.56	52.04	01.59	02.79
• 17	0.5661	-0.3929	47.95	08.66	03.10	40.79	01.10	54.99
• 18	1.1406	-0.4850	81.28	32.78	30.14	53.34	03.68	12.82
• 19	1.6009	0.2997	65.06	95.03	61.61	02.96	32.94	02.46
• 20	1.4832	-0.4491	96.61	51.70	51.09	45.52	00.88	02.49

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 [36] - see also [34] - 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

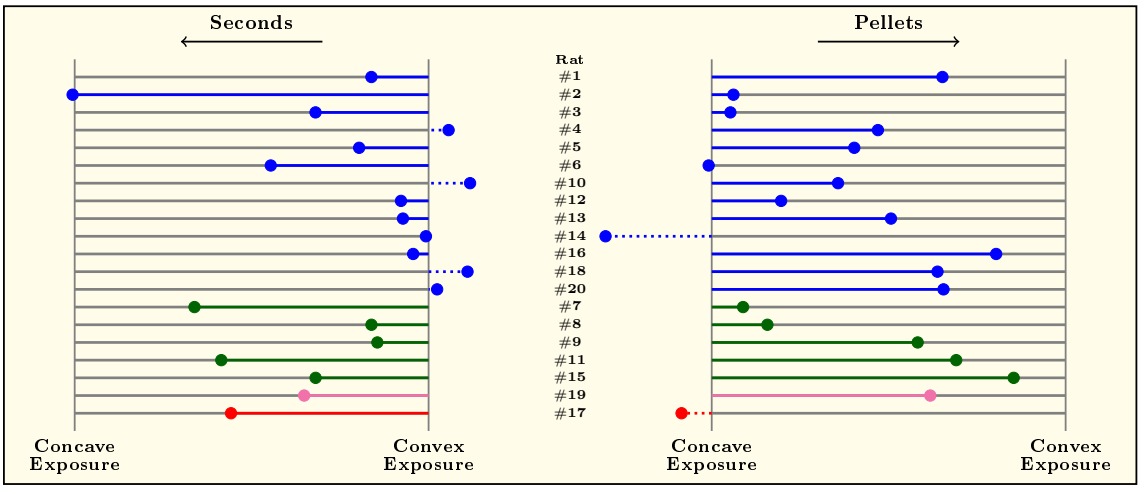
				Rewards (pellets)					Losses (seconds)				
Rat	Nosepokes	Black Swan	Jackpot	Sum	Mean	Var.	Sk.	Kur.	Sum	Mean	Var.	Sk.	Kur.
• 1	5761	3	13	8399	3.08	33.05	11.54	152.99	27408	9.04	67.54	22.12	618.13
• 2	5526	13	8	7622	2.92	23.13	12.88	203.11	28305	9.69	274.50	12.08	166.54
• 3	5006	6	4	6896	2.92	12.48	17.82	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. 23

[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. 19, 20

[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. 7

[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. 21

[8] Chevin L.-M., and Hoffmann A.A. (2017). Evolution of phenotypic plasticity in extreme environments. Phil. Trans. R. Soc. B 372, 20160138. 21, 22

[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. 3, 21

[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. 3, 20, 32

[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. 21

[12] Denrell, J., March, J. G. (2001). Adaptation as information restriction: The hot stove effect. Organization Science, 12(5), 523-538. 20

[13] 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. 22

[14] 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, 319-351. 20

[15] 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. 21

[16] 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. 30, 32

[17] Fletcher, R., Reeves, C. M. (1964). Function minimization by conjugate gradients. Computer Journal 7, 148-154. 37

[18] 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. 20

[19] 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

[20] Hertwig, R., and Erev, I. (2009). The description-experience gap in risky choice. Trends in Cognitive Sciences, 13: 517-523. 3

[21] 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

[22] Kahneman D., and Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica 47, 263 20

[23] 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

[24] 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

[25] Loehman, E. (1994). Rank dependent expected utility: stochastic dominance, risk preference, and certainty equivalence. Journal of Mathematical Psychology, 38: 159-197. 32

[26] 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. 21

[27] Mason, A., Ludvig, E.A., Spetch, M.L., Madan, C.R. (2024). Rare and extreme outcomes in risky choice. Psychonomic Bulletin & Review, 31: 1301-1308. 3, 22

[28] 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

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

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

[31] Nelder, J., Mead, R. (1965). A simplex method for function minimization, Computer Journal, vol. 7(4), 1965, 308-313. 37

[32] Newell, B., Rakow, T., Yechiam, E. et al. (2016). Rare disaster information can increase risk-taking. Nature Clim Change 6:158-161. 3

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

[34] Palminteri, S., Lebreton, M. (2021). Context-dependent outcome encoding in human reinforcement learning. Current Opinion in Behavioral Sciences: 41: 144-151. 47

[35] Palminteri, S., Wyart, V., Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. Trends in Cognitive Sciences: 21(6), 425-433. 38

[36] Parducci, A. (1995). Happiness, Pleasure, and Judgment: The Contextual Theory and its Applications. Lawrence Erlbaum Associates, Inc. 47

[37] 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

[38] Plonsky O, Teodorescu K, Erev I. Reliance on small samples, the wavy recency effect, and similarity-based learning. Psychol Rev. (2015) 122:621-47. 20

[39] Prelec, D. (1998). The Probability weighting function. Econometrica, 66:497-527. 32

[40] Real, L.A (1980). Fitness, uncertainty, and the role of diversification in evolution and behavior. The American Naturalist, 115: 623-638. 7

[41] Real, L.A. (1991). Animal choice behavior and the evolution of cognitive architecture. Science, 253: 980-986. 3

[42] Redish A.D., et al. (2021). Computational validity: using computation to translate behaviours across species. Phil.Trans. R. Soc., B377: 20200525. 21

[43] 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. 7

[44] 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 20

[45] Rushworth M.F., Behrens T.E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. Nature Neuroscience. 11(4):389-97. 21

[46] 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. 21

[47] 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

[48] Simons A. M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. Proceedings. Biological sciences, 278: 1601-1609. 7

[49] 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

[50] Sutton, R.S., Barto, A.G. (2020). Reinforcement learning: An introduction (2nd ed.). The MIT Press. 14

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

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

[53] 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. 3

[54] 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. 20, 21

[55] 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. 21

[56] 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. 21

[57] 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

[58] van Wingerden M., Marx C., Kalenscher T. (2015). Budget constraints affect male rats' choices between differently priced commodities. PLoS ONE: 10(6). 3

[59] 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. 19

[60] Yamamori Y., Robinson O.J., Roiser J.P. (2023). Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance. Elife Nov 14:12:RP87720. 20

[61] 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): 232–43 20, 21

Point by point replies to the Reviews (that appear in blue)

Dear Reviewers,

We thank you very much for the constructive and helpful comments that appear in the reviews. We hope you will accept our sincerest apologies for the delay in delivering a revised version, which has been entirely rewritten.

We provide below point by point replies (in blue) to all the weaknesses and recommendations stressed by both Reviewers, indicating how and where in the revised manuscript we have addressed them. To help the reading, we have indicated in blue in the revised version the changes made in the revised draft, which we hope satisfactorily address all the points raised by the Editors and Reviewers.

We hope that the revised version meets the journal's standards and we would like to thank you again for allowing us on a path that we think has led to a much improved manuscript.

Best wishes,

Patrick Pintus, on behalf of all co-authors
Professor of Economics
Aix-Marseille University

Public Reviews:

Reviewer #1 (Public Review):

Summary:

In this manuscript, the authors investigate the impact of rare and extreme events on rodents' decision-making under risk, in gain and loss contexts. They describe the behavior of 20 rats performing a

four-armed bandit task, where probabilistic gains (sugar pellets) and losses (time-out punishments) can - in some arms - incorporate extremely large - but rare - outcomes. They report that most rats are sensitive to rare and extreme outcomes despite their infrequent occurrence, and that this sensitivity is primarily driven by extreme loss events which they try to avoid, rather than extreme gains that they seek to obtain.

They finally propose a modification of standard reinforcement-learning, which features a specific sensitivity to rare and extreme outcomes and can account for the observed behavior.

Strengths:

The manuscript really taps into a surprisingly neglected but very relevant aspect of decision-making: the effect of rare and extreme events (REE). The authors have developed an experimental setup that seemingly allows investigation of this aspect, which is not trivial given the idiosyncratic properties of rare and extreme events.

The parameters of the experimental setup seem also to be well thought off: basically, in the absence of REE, some options are objectively better than others (because, in expectation, they overall deliver more food, or minimize time-out punishments), but this ordering reverses if REE are taken into account. This allows for a clean test of the integration of REE in the rodent's decision-making model.

The data is presented and analyzed in a very descriptive but exhaustive and transparent way, down to the description of individual rodent's behavior.

Weaknesses:

While the description and analyses of the behavioral patterns are rigorously done under the economic lens of risky decision-making, the authors' interpretation heavily relies on the assumption that rodents have built the correct model of the task during the training. Extensive details are provided about the training procedure, and the observed behavior at the end of the training, but it remains virtually impossible to disambiguate choices due to imperfect learning to choices made due to intrinsic preferences for risk or REE.

Reply: As detailed in Material and Methods, the animals were progressively overtrained following standard behavioral procedures. During this process, they experienced all available options, including both positive and negative REE. We assume that repeated exposure to these REE supported learning, as would be expected for any event occurring throughout such an extended training phase. The rats ultimately displayed an asymmetric pattern of choices: they consistently avoided the Black Swan, indicating that they had learned its negative consequences, yet they did not systematically seek the Jackpot. If their behavior were driven solely by incomplete learning or by an inherent preference for

risk or REE, we would expect to see the opposite pattern—systematic Jackpot seeking or inconsistent avoidance of the Black Swan.

By nature, gains (food pellets) and losses (time-out punishments) are somewhat incommensurable so the interpretation of the asymmetry due to outcome valence is also subject to interpretation. There might be some additional subtleties due e.g. satiety that could come from gaining REE (i.e. the delivery of 80 pellets from the Jackpot).

Reply: As described in Material and Methods, we used mouse pellets (20 mg) instead of rat pellets (45 mg) to prevent satiety during Jackpot delivery (80 pellets). We also selected gains (sweet pellets) and losses (delays) that we have successfully used in previous rat decision-making paradigms, such as the rat gambling task (Adams et al., 2017; doi: 10.1523/ENEURO.0094-17) and the loss-chasing task (Breysse et al., 2021; doi: 10.1111/ejn.14895). Notably, if the Jackpot induced satiety, one would expect animals to stop seeking it—yet this was not systematically observed. Nonetheless, we added a sentence to the Discussion on page 18 of the manuscript to acknowledge that we cannot fully exclude the possibility that satiety contributed to the lack of systematic Jackpot Seeking.

In its current form, the paper is quite hard to digest. This is naturally the case with interdisciplinary work (here mixing economists and neurobiologists). But I am afraid that with the current frame, the paper is going to miss its target, in terms of audience.

Reply: We have rewritten entirely and the english was corrected thanks to ChatGPT. We hope that the paper is now easier to digest.

The proposed model seems somewhat disconnected from the behavioral patterns: while the model suggests an effect of REE at the decision stage (i.e. with specific decision weights for those rare events), this formalism seems at odds with the observation that REE (notably in the loss domain) has an impact of subsequent behavior - (Black Swans tend to reinforce Total Sensitivity to REE) which rather suggests an effect at the learning stage.

Reply: We agree with the referee that this may appear surprising at first glance. However, we would first like to emphasize that the general model allows REE to influence learning—that is, to contribute to the updating of the Q-subvalues. Moreover, even when REE are incorporated only as decision weights, as is the case for most rats, this does not imply that REE are unimportant during learning. In fact, the model assumes that REE are learned once and for all when they first occur during a trial of the corresponding option. Unreported simulation exercises indicate that a more gradual learning of maximal and minimal values would likely yield similar results.

Second, the Before/After analysis shows that the behavioral response to Black Swans is locally small in terms of both total and one-sided sensitivities. This suggests that such effects are likely too subtle to be captured by this class of models for most rats. We have added this clarification to the revised version (page 17).

Discussion:

This study convincingly demonstrates that REEs are processed rather uniquely, which makes sense given their evolutionary relevance. REE has indeed been somewhat neglected in previous research, and this study therefore opens an interesting new front on the fundamental aspects of decision under

risk. The authors have devised an original theoretical and empirical framework that will be useful for the community, and the combination of economics analysis and rodent behavior constitutes a thought-provoking ground to think about the nature of risk preferences. The interpretation and mechanistic account of these aspects, as well as their generalizability outside the specific context of this study, remain to be strengthened.

Reply: We have modified the discussion to further insist on the translational aspect of the study and its interest for various populations (page 22). We hope that the generalizability is now strengthened.

Reviewer #2 (Public Review):

Summary:

This paper attempts to examine how rare, extreme events impact decision-making in rats. The paper used an extensive behavioural study with rats to evaluate how the probability and magnitude of outcomes impact preference. The paper, however, provides limited evidence for the conclusions because the design did not allow for the isolation of the rare, extreme events in choice. There are many confounding factors, including the outcome variance and presence of less-rare, and less-extreme outcomes in the same conditions.

Strengths:

(1) The major strength of the paper is the significant volume of behavioural data with a reasonable sample size of 20 rats.

(2) The paper attempts to examine losses with rats (a notoriously tricky problem with non-human animals) by substituting time-outs as a proxy for losses. This allows for mixed gambles that have both gain and loss possible outcomes.

(3) The paper integrates both a behavioural and a modelling approach to get at the factors that drive decision-making.

(4) The paper takes seriously the question of what it means for an event to be rare, pushing to less frequent outcomes than usually used with non-human animals.

Weaknesses:

(1) The primary issue with this work is that the primary experimental manipulation fails to isolate the rare, extreme events in choice. As I understand the task, in all the conditions with a rare extreme event (e.g., 80 pellets with probability epsilon), there is also a less-rare, less-extreme event (e.g., 12 pellets with probability 5). In addition, the variance differs between the two conditions. So, any impact attributable to the rare, extreme event could be due to the less rare event or due difference in the variance. The design does not support the conclusions. Finally, by deliberately confounding rarity and extremity, the design does not allow for assessing the impact of either aspect.

Reply: We agree with the referee that both the REE and the rare (≈10% frequency) but non-extreme outcomes are present in the relevant options. However, the rare but non-extreme reward is not large

enough to make the convex option attractive and to shift choice away from the concave option. In other words, unlike REE, these outcomes do not reverse stochastic dominance in our design (as noted in *Material and Methods*). We have explored modified designs for human subjects in which the rare but non-extreme outcomes are removed. Preliminary results indicate that the behavioral phenotypes observed in rats also emerge in humans under these modified conditions, suggesting that REE are the primary drivers. We have added a statement to the Discussion (page 22) to clarify this point.

We elaborate further in our response to point (3) below on why analyses based solely on variance are insufficient when dealing with REE. To clarify the role of rare and extreme outcomes in distinguishing convex from concave options, we provide two new columns to Table 2 in the *Materials and Methods*, in our reply to point (3).

Finally, although a detailed analysis of rare but non-extreme outcomes lies outside the scope of this paper, the symmetric treatment of extreme and frequent outcomes can be addressed straightforwardly using strong First-Order Stochastic Dominance. Classical decision-theoretic approaches indeed satisfy this property.

(2) The RL-modelling work also fails to show a specific impact of the rare extreme event. As best as I can understand Eq 2, the model provides a free parameter that adds a bonus to the value of either the two options with high-variance gains (A and V in the paper) or to the two options with high-variance losses (F and V in the paper). This parameter only depends on whether this option could have possibly yielded the rare, extreme outcome (i.e., based on the generative probability) and was not connected to its actual appearance. That makes it a free parameter that just bumps up (or down) the probability of selecting a pair of options. In the case of the "black swan" or high-variance loss conditions, this seems very much like a loss aversion parameter, but an additive one instead of a multiplicative one.

Reply: We agree with the referee that the additional parameters, compared to more standard Q-learning models, specifically capture the fact that some options deliver REE while others do not. In our estimation procedure, these parameters become nonzero as soon as REE are observed for the first time for a given option. Therefore, the first step is to estimate a baseline nested model in which REEs contribute only at the learning stage (i.e., they affect the updating of Q-subvalues), while the additional parameters are constrained to zero. The next step is to compare alternative models against this baseline, allowing REEs to enter through the additional parameters. In this respect, our specification is parsimonious, especially given that very little is known about REEs in computational neuroscience. More structural modeling is certainly a promising direction for future research, and this paper constitutes a first step toward that goal.

We provide the BIC, in addition to the AIC, to account for the presence of additional parameters in model selection and to ensure that the observed improvement in fit is not merely driven by their inclusion.

Unlike most of the existing literature, our results extend the notion of loss aversion to **extreme** losses. The negative decision weight on options yielding the Black Swan can be interpreted as a differential treatment of negative REE, an issue we discuss extensively in the Discussion (page 20).

(3) The paper presented the methods and results with lots of neologisms and fairly obscure jargon (e.g., fragility, total REE sensitivity). That made it very hard to decipher exactly what was done and what was found. For example, on p. 4, the use of concave and convex was very hard to decipher; the text even has to repeat itself 3 times (i.e., "to repeat" and "in other words") and is still not clear. It would be much clearer (and probably accurate) to say that the options varied along the variance dimension, separately for gains and losses. Option A was low-variance gains and losses. Option B was low-variance losses and high-variance gains. Option C was high-variance losses and low-variance gains, and Option D was high-variance losses and gains. That tells much more clearly what the animals experienced without the reader having to master a set of new terminologies around fragility and robustness, which brings a set of theoretical assumptions unnecessarily into the description of the experimental design. In terms of results, "Black Swan" avoidance is more simply known as risk aversion for losses.

Reply: Because our experimental design focuses on REE, outcomes cannot be summarized only by their variance. This is well known from the large literature on so-called fat-tailed statistical distributions. Unlike the Normal distribution that is entirely characterized by its expected value and variance, fat-tailed distributions have nonzero kurtosis. This implies that a fat-tailed distribution (e.g. exponential) with the same expected value and variance as the Normal differs importantly by possessing extreme values that are much more likely in terms of frequency. To illustrate, if the distribution of pellets was assumed to be Normal with expected value set at 3.89 and variance set at 9.37 as for the convex option, the probability of getting 80 pellets would be about 2.10^{-16}, practically zero. In contrast, this probability is smaller than, but close to 1% in our design.

In Material and Methods, we clearly explain how our novel approach in terms of convexity relates to the moments of the reward distributions, including but not limited to the variance. To clarify further, we provide below two new tables (Tables A and B) to be compared to Table 2 of the manuscript in which we report the first four moments (mean, standard deviation, skewness and kurtosis) of the full concave and convex gain distributions, reproduced below for convenience

	Convex NE+RE+REE	Concave NE+RE+REE
Mean	3.89556	3.11778
Std Deviation	9.37942	1.14577
Skewness	7.16936	0.203185
Kurtosis	57.8507	1.38091

In Table A we report the first four moments when REE are truncated. Comparing convex and concave gains shows that the convex option has a smaller but still close mean compared to the concave option. In contrast, the former has larger variance, skewness and kurtosis compared to the latter. Therefore, interpreting choosing the convex option as reflecting “preference” for variance is at best incomplete.

Table A: first four moments of concave and convex gains when REE are removed

	Convex NE+RE	Concave NE+RE
Mean	2.86712	3.09234
Std Deviation	3.13623	1.13226
Skewness	2.28563	0.218717
Kurtosis	7.02751	1.38101

Table A further shows that REE alone goes a long way towards explaining the differences between convex and concave options in terms of the first four moments: removing the rare and extreme value results in the concave option having now a **larger** mean, while the convex option still has larger variance, skewness, and kurtosis but by a smaller margin.

In Table B we report the first four moments when both RE and REE are truncated, which shows that the convex and concave options differ only with respect to their mean (which is here also larger for concave).

Table B: first four moments of concave and convex gains when both RE and REE are removed

	Convex NE	Concave NE
Mean	1.88778	2.88778
Std Deviation	0.993683	0.993683
Skewness	0.225866	0.225866
Kurtosis	1.05102	1.05102

In addition, our focus on REE implies that we go beyond mean-variance preferences that apply mostly to Gaussian distributions. It is not clear theoretically what type of utility functions would reflect preferences that combine a taste for variance, skewness and kurtosis, even though all those moments affect expected utility. See for example Phelps, C.E. "A user's guide to economic utility functions". J Risk Uncertain 69, 235–280 (2024) for a recent overview (on page 242, Phelps states that "In situations where risk is not normally distributed, it is ill-advised to ignore statistical parameters beyond variance, unless the deviations from normality are relatively small").

More importantly, our proposed measure of the convexity of the reward distributions, the Jensen gap, further reveals how even restricting the analysis to the first four moments is incomplete in the sense that it fails to characterize the difference between options: the fifth moment of the concave contributes more the Jensen gap than even kurtosis, while one needs to look at much higher moments to find significant contributions to the Jensen gap for the convex option. In that sense, there is no reason to restrict the analysis to variance, and even to skewness and kurtosis, to compare options, in general and in our particular setup as well. Note that introducing REE would result in convex distributions even in simplified designs, e.g. with 3-value support. Studying REE implies the need to look beyond variance, and our proposal is to use the Jensen gap as a measure of convexity. In the Material and Methods section of the paper, we did not develop an in depth analysis of Jensen gap so as to spare the reader confronted with an already rather technical paper.

We thank the referee for raising the issue of whether variance is a simpler explanation of our results. To keep the main text as short as possible, we chose to refrain from adding technical complexity. We hope we made clear in our reply that the analysis cannot be restricted to variance when studying REE. We believe that Jensen gap is a useful notion in this regard. As our replies will be made publicly available, we chose not to integrate the above discussion in the main text.

(4) Were the probabilities shuffled or truly random (seem to be fixed sequences, so neither)? What were the experienced probabilities? Given the fixed sequences, these experienced ("ex-post") probabilities, could differ tremendously from the scheduled ("ex ante") probabilities. It's quite possible that an animal never experienced the rare, extreme event for a specific option. It's even possible (if they only picked it on the 10th/60th choices by chance), that they only ever experienced that rare extreme event. This cannot be known given the information provided. The Supplemental info on p.55 only gives gross overall numbers but does not indicate what the rats experienced for each choice/option-which is what matters here. A simple table that indicates for each of the 4 options, how often they were selected, and how often the animals experienced each of the 6-8 possible outcome would make it much clearer how closely the experience matched the planned outcomes. In addition, by restricting the rare outcome to either the 10th or 60th activations in a session, these are not random. Did the animals learn this association?

Reply: Probabilities are not random and a limited number of fixed sequences has been used, as stated in Material and Methods. We have chosen sequences that satisfy our assumptions about ex-post stochastic dominance reversal of convex over concave options when REE are added. We have added in Table S4 the choice frequencies for all four options. If the animals had learnt the 10th and 60th activation, they would exhibit a strategy in their choice that would tend to be more optimized than what is observed. For example, the options offering the possibility to obtain the Jackpot are not optimal in terms of gains for the frequent events, therefore the animals should tend to select these options only around the 10th and 60th choice. Most of their other choices should favor the options delivering the larger gains in the frequent domain. This is not what is observed. We have added this important point in the discussion (page 18).

(5) The choice data are only presented in an overprocessed fashion with a sum and a difference (in both figures and tables). The basic datum (probability/frequency of selecting each of the 4 options) is not provided directly, even if it can theoretically be inferred from the sum and the difference. To understand what the rats actually do, we first need to see how often they select each option, without these transformations.

Reply: As described in Material and Methods, the 4 options are combinations of 2 convex and concave sub-options for gains and losses, which is why our analysis of the behavioral data focuses on convexity-related total and one-sided sensitivities to REE. The third dimension needed to fully characterize rats' behavior is simply $1 - f_F$, the fraction of non-Fragile choices. In addition, we also provide in Table S4 of the Supplementary Material an alternative interpretation in terms of Black Swan Avoidance and Jackpot Seeking. We have added in Table S4 the choice frequencies for all four options. Finally, all the raw data will be made available with open access and no access codes.

(6) There is insufficient detail provided on the inferential statistical tests (e.g., no degrees of freedom or effect sizes), and only limited information on exactly what tests were run and how (bootstrapping, but little detail). Without code or data (only summary information is provided in the supplement), this

is difficult to evaluate. In addition, the studies seem not to be pre-registered in any way, leaving many researchers with degrees of freedom. Were any alternative analysis pipelines attempted? Similarly, there were many sub-groupings of the animals, and then comparisons between them - were these post-hoc?

Reply: We understand the concern of the referee for pre-registration of the referee, as an epistemic safeguard to make empirical claims more falsifiable, more transparent, and less dependent on post hoc rationalization. But the contemporary push for preregistration is often presented as an "epistemic improvement," but in practice it functions largely as a norm of moral regulation, not a scientific necessity. The rhetoric is moralistic: preregistered research is "clean," "transparent," "credible," while non-preregistered work is viewed with suspicion—even when the methodology is sound. This language is not epistemologically neutral; it enforces *ought* to be done, irrespective of the diversity of legitimate scientific practices.

From a philosophy of science perspective, this is historically and conceptually problematic. Scientific progress has never followed a uniform, rule-based method. As e.g. Feyerabend has argued, major discoveries have emerged precisely because researchers were **not** bound by predetermined plans: they followed anomalies, improvised, reinterpreted data, and revised methods and hypotheses in light of new evidence — practices that a rigid preregistration ethos can suppress and that are not aligned with how genuine discovery often occurs.

Even from a statistical standpoint, preregistration is far from a panacea. It reduces some degrees of freedom (mainly in confirmatory statistics), but it does not eliminate flexibility; researchers can still choose models, transformations, exclusion rules, stopping rules, etc. And more importantly: reducing flexibility is not inherently epistemically virtuous. Flexibility is often *necessary* to understand data properly—especially in new paradigms or first-of-their-kind experiments, which is the case for this study. Science needs exploration, opportunism, and theoretical plasticity. Preregistration is compatible with these only if it is treated as one optional tool among many—not as a universal evaluative standard.

As the referee pointed out, this study "taps into a surprisingly neglected but very relevant aspect of decision-making." Our work is therefore mainly exploratory: the experimental paradigm reveals new behavioral patterns in how rats cope with rare and extreme events, and much of our analysis is necessarily descriptive. We conduct formal inference only where it is methodologically appropriate — the short-term behavioral response to rare events (for which we now provide more details in the Material & methods section p.35) and the estimation of augmented Q-learning models, which follow a standard econometric approach (documented in the Material & Method section–see also our response to recommendation 4). These inferential results support the descriptive patterns that motivate this new line of research.

(7) On p. 17, there is an attempt to look at the impact of a rare, extreme event by plotting a measure of preference for the 10 trials before/after the rare, extreme event. In the human literature, the main impact of experiencing a rare, extreme event is what is known as the wavy recency effect (See Plonsky et al. 2015 in Psych Review for example). What this means is that there tends to be some immediate negative recency (e.g., avoiding a rare gain) followed by positive recency (e.g., chasing the rare gain).

Using a 10-trial window would thus obscure any impact of this rare, extreme event. An analysis that looks at a time course trial-by-trial could reveal any impact.

Reply: We thank the referee for drawing our attention to the wavy recency effect documented in human experiments. We have added the corresponding reference in the Discussion (page 20). Regarding rats, the Before/After analysis reported in the paper suggests that there is no sizeable immediate recency effect for Jackpots. Even for Black Swans, the immediate recency effect we report remains modest when using a 10-trial window, and the analysis of the choice immediately following a REE does not show evidence of immediate negative recency. This casts doubt on the presence of such an effect in rats.

(8) As I understood the method (p. 31), the assignment of options to physical locations was not random or counterbalanced, but deliberately biased to have one of the options in the preferred location. This would seem to create a bias towards a particular option and a bias away from the other options, which confounds the preference data in subsequent analyses.

Reply: We agree that the design incorporated an intentional bias toward the antifragile option as a proof of concept. Nevertheless, Figure 8 demonstrates that animals substantially altered their choices between training and final testing, with a median change of approximately 35% across sessions. This indicates that behavior was driven by the structure of possible outcomes rather than by a stereotyped location-based preference.

(9) Are delays really losses? This is a big assumption. Magnitude and delay are different aspects of experience, which are not necessarily commensurable and can be manipulated independently. And, for the model, how were these delays transformed into outcomes for the model? Eq 1 skips over that. Is there an assumption of linearity? In addition, I was not wholly clear if the delays meant fewer trials in a session or if the delays merely extended the session and meant longer delays until the next choice period.

Reply: Consistent with established rodent decision-making paradigms (Adams et al., 2017 doi: 10.1523/ENEURO.0094-17; Breysse et al., 2021 doi: 10.1111/ejn.14895), we employed sweet pellets as gains and imposed delays as losses. Delays are operationalized as losses because they preclude the animal from engaging in reward-generating behavior; thus, increasing the delay duration proportionally increases the magnitude of the opportunity cost.

(10) The paper does not sufficiently accurately represent the existing literature on human risky decision-making (with and without rare events). Here are a few examples of misrepresented and/or missing literature:

-Most studies on decision-making do not only rely on p > 10% (as per p. 2). Maybe that is true with animals, but not a fair statement generally. Some do, and some don't. There is substantial literature looking at rarer events in both descriptions (most famously with Kahneman & Tversky's work), but also in experience (which is alluded to in reference 19). That reference is not only about the situation when choices are not repeated (e.g. the sampling paradigm), but also partial feedback and full-feedback situations.

Reply: We have corrected that statement in the main text (page 3) and we thank the referee for pointing this out.

The literature on learning from rewarding experiences in humans is obliquely referenced but not really incorporated. In short, there are two main findings - firstly people underweight rare events in experience; second, people overweight extreme outcomes in experience (both contrary to description). Some related papers are cited, but their content is not used or incorporated into the logic of the manuscript.

One recent study systematically examined rarity and extremity in human risky decision-making, which seems very relevant here: Mason et al. (2024). Rare and extreme outcomes in risky choice. Psychonomic Bulletin & Review, 31, 1301-1308.

There is a fair bit of research on the human perception of the risk of rare events (including from experience) and important events like climate. One notable paper is Newell et al (2015) in Nature Climate Change.

Reply: We agree with the referee that the related literature on REE in animal Decision Making is scant and that it is more developed in humans. We thank the referee for pointing at Mason et al. (2024), who clarify where the literature on humans stands and why combining rarity and extremity, as we also do, is important and highly relevant. We have added a new statement and references in the Introduction and Discussion (pages 3, 20, 22) .

Recommendations for the authors: please note that you control which revisions to undertake from the public reviews and recommendations for the authors.

Reviewer #1 (Recommendations For The Authors):

(1) As said above, I think the manuscript would really benefit from a rewriting, to replace some technical terms with more readable ones, and maybe rebalance the focus from the current focus on the framework (heavily loaded with economics concepts, which will be hard to digest for the eLife readership) to a higher weight on information that is critical to understand and interpret the behavior (e.g. information about training & training behavior, etc.).

Reply: We have revised the entire manuscript to improve readability and have clarified in the main text: (1) why convexity of exposures to REE could, beyond variance, be useful for experiments in other settings that our own; (2) why the associated notion of antifragility may be applicable to other settings and therefore of broader interest; (3) what was done in the training sessions compared to the final sessions.

(2) From Figure 8, it seems that rodent behavior is more clustered after the training (i.e. before the sessions) than after the sessions. Could that be a sign of imperfect learning?

Reply: Figure 8 mostly suggests that there is some flexibility in the choices made and that the intended initial bias towards the antifragile choice in the design of the task could be over ridden by the rats.

(3) The modelling section seems incomplete. I think the authors want to tease apart where REE enters the model and should propose an alternative where REE affects the learning rather than the decision.

Reply: In fact, the general model allows REE to have an effect at the learning stage only (i.e. to contribute to the updating of the Q subvalues), when the specific decision weights attached to options delivering REE are both zero. However, our analysis shows that such a model is rejected by the behavioral data for all rats. We have clarified this point in the revised version.

(4) Also, parameter and model recovery exercises seem mandatory (Wilson & Collins, 2019).

Reply: We thank the referee for highlighting this valuable reference in computational modeling, particularly in the context of model identification and estimation in computational biology. In the present research, we adopted an econometric perspective on model identification—especially with regard to the integration of Q-values for gains and losses. The softmax choice function is formally equivalent to a multinomial logit model, and as is well known in econometrics, identification in such models presents non-trivial challenges. The standard approach in classical Q-learning is to multiply the Q-value by an inverse temperature parameter (also known as a precision parameter in random utility models). When extending the model to include separate Q-values for gains and losses, specifying the model in an identifiable way becomes more complex.

To address this issue, we considered several alternative model specifications and conducted grid-based estimation of starting parameter values. This approach allowed us to examine the shape of the log-likelihood function and assess whether the parameters are globally identified, rather than only identifiable up to a linear combination. We found that the most parsimonious and empirically identified specification in our experimental paradigm is one in which Q-values for gains and losses are summed, each weighted by distinct decision weights (see our Equation 2 in the paper).

The inclusion of decision weights for REE for each option (Equation 2) is then structurally equivalent to introducing constant terms in a logit model. The identification of these parameters follows standard econometric results on discrete choice models (e.g., Davidson & MacKinnon, 2003): since we model choices among four options, three free parameters can be estimated, leaving one degree of freedom in the specification. As mentioned in the "Modelling and Statistical Analysis" section, we further guarded against the presence of local maxima by applying a two-step estimation procedure, combining two optimization algorithms with multiple sets of starting values for the baseline model (i.e., the model without decision weights for REE). We also tested the addition of a global optimization method—simulated annealing—but found that it did not significantly improve upon our two-step procedure. This is not surprising, as our preliminary investigation of model identification, based on grid searches over starting parameter values, confirmed that all parameters were identified in our simple specification. Our intuition is that simulated annealing may yield different estimates than gradient-based methods primarily in cases where the model is not theoretically identified—suggesting that the need for such global optimization techniques can be indicative of underlying identification issues in Q-learning models.

Regarding model comparison, we have used penalized information criteria to account for additional parameters. Although we do not report confusion or inversion matrices for our nested models, we verified that the estimated models replicate observed behaviors across all phenotypes, as shown in the main text (see bottom left panel of Figure 5 for the Total and One-Sided sensitivities). Most importantly, we conducted 100 additional simulations of 40 artificial sessions for each phenotype using the “winning” models and the median fitted parameters. These simulated rats—playing the task 100

times over 40 sessions—offer strong evidence that the selected models are valid: they quantitatively capture the behavior of all phenotypes in terms of our key metrics, Total and One-Sided sensitivities (see bottom right panel of Figure 5).

Taken together, this methodical econometric approach to model specification and estimation gives us strong confidence in the identification and robustness of our model. Overall, while Wilson & Collins (2019) provide an interesting framework for model estimation in computational biology, we believe that a more formal theoretical analysis of model identification in Q-learning models would be a valuable addition to the field—though it lies beyond the scope of the present work. In our view, computational biologists should complement simulation-based validation and empirical fit with formal methods for assessing theoretical identifiability, particularly when estimating complex choice models.

Davidson, R. and J.G. MacKinnon (2003) *Econometric Theory and Methods. Oxford University Press (New York).*

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. eLife, 8, e49547. https://doi.org/10.7554/eLife.49547

Reviewer #2 (Recommendations For The Authors):

(1) The paper confuses risk sensitivity and exploration in the opening lines. These are not the same.
Reply: What we have in mind here is that uncertainty about outcomes is one of the main drivers of exploration, in the sense that there would be no need to explore in a counterfactual world with deterministic gains and losses. We have modified the opening lines of the paper to better reflect this dimension (page 2).

(2) p. 9. "awfully long" is an unnecessary descriptor. Descriptions of methods should be more factual.
Reply: The manuscript has been entirely rewritten.

(3) p. 13. Most points lie on the left of the square (not right?).
Reply: We thank the referee for pointing at this typo, that is now corrected in the text (page 8).

(4) p. 13. Last line. "obviously" is patronizing to the readers.
Reply: The manuscript has been entirely modified to address related points.

(5) p. 23. The avoidance of black swans by not choosing that option sounds like a hot-stove effect (see Denrell & March, 2001). Is this evidenced here?
Reply: To the best of our knowledge, the statement that "people tend to avoid activities they have had a negative experience of, resulting in a negativity bias" (from Jerker Denrell's website) does not explicitly concern REE. Instead, it appears to refer broadly to reinforcement learning mechanisms driven by negative outcomes, irrespective of their magnitude or frequency. In our task, animals encounter both negative rare events (RE) and negative rare and extreme events (REE; Black Swans). Notably, the task design does not allow rats to completely avoid negative RE unless they cease performing the task altogether—a pattern typically seen in paradigms involving aversive stimuli such as electric foot shocks. The fact that all 20 rats maintained stable performance across the 41 sessions provides evidence against a pronounced hot-stove effect. This point has been incorporated into the revised discussion (page 20).

(6) "menus" is an odd term. Better described as reward schedules?
Reply: “Menu” has been replaced by “option” in the main text.

(7) Why are they 20-minute sessions? I thought it was 120 trials per session? And 41 sessions? Or was this only in training?
Reply: Each session ended after 20 minutes had elapsed, which led to approximately 120 trials (but not systematically). The choice of 20 minutes was made in order to limit the number of trials to prevent satiety. The total number of sessions ran with all 20 animals for the final testing was 41, an odd number but there was no justification to remove one session from the analysis. The training was much longer and is not included in the 41 sessions.

(8) Really not clear why these Jensen inequalities were relevant or even calculated for these options? How is it relevant to what animals chose or experienced? They seem to be based on the generative probabilities for different options, which is not what happened in reality.
Reply: We propose the Jensen gap as a general measure of convexity that relates to all moments of the probability distribution, as described in more detail in our answer to point (3) above. As such, we think it is a characterization of options with stochastic outcomes that could prove useful to other experimenters in alternative settings beyond our own.

(9) Only some summary data in supplemental materials. No open data or code for recreating the experiment or analyzing the data.
The data is available on Github (see page 38) and the code will be available upon request.

95 Regent Street
Cambridge
CB2 1AW, UK

P 01223 855340
W elifesciences.org
X @eLife

January 5, 2026

Publication in eLife

To whom it may concern:

We are pleased to confirm that eLife has reviewed the following paper, which has been published as a Reviewed Preprint:

Degoulet M, Willem L-M, Baunez C, Luchini S, Pintus PA (2024) Specific Sensitivity to Rare and Extreme Events: Quasi-Complete Black Swan Avoidance vs Partial Jackpot Seeking in Rat Decision-Making. *eLife* **13**:RP98487. DOI: https://doi.org/10.7554/eLife.98487

Editors: Michael J Frank, Brown University, United States; Angela Langdon, National Institute of Mental Health, United States

The revised version was received on December 5, 2025, and a revised version of the Reviewed Preprint will be published online in the coming weeks. The version eLife is currently reviewing is available on bioRxiv (https://www.biorxiv.org/content/10.1101/2021.11.01.466806v4).

eLife is an independent non-profit committed to improving the way research is reviewed and communicated. A Reviewed Preprint includes the preprint itself, the eLife Assessment, Public Reviews that describe the strengths and weaknesses of the article, and a response from the authors (if available). The authors can choose to publish their article as a regular journal article (the Version of Record), usually after a round of revisions, to mark the end of the review and publication process. A Version of Record complies with all major funding agency requirements for immediate online access to the published results of their research grants.

Best wishes,

A Collings

--
Andy Collings
Executive Editor, *eLife*