



### **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 (Breyse 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 ( $\approx 10\%$  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 than 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 overridden 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.