When jackpot misleads: The disrupting role of rare rewards in value learning

Luning He (luninghepsy@gmail.com)
School of Psychological Science, University of Bristol,
Bristol, United Kingdom

Anina N. Rich (anina.rich@mq.edu.au)
School of Psychological Sciences, Macquarie University,
Sydney, Australia

Yinan Cao (ycaoneuro@gmail.com)
Département d'Études Cognitives, Ecole Normale Supérieure,
Paris, France

Konstantinos Tsetsos (k.tsetsos@bristol.ac.uk)
School of Psychological Science, University of Bristol,
Bristol, United Kingdom

Abstract

Prior research suggests that, when making decisions from experience, people tend to undervalue high-risk, high-reward options. This tendency has been attributed to underweighting the impact of rare rewards. However, this underweighting view may be confounded with risk avoidance. Here, we challenged a pure "underweighting" account using a two-armed bandit task. We found that choices for options with large but rare rewards were insensitive to their expected values. Instead, they were guided by a value-independent sampling bias for the rare-reward option. These findings suggest that the presence of large but rare rewards disrupts value-based decision-making, shifting the decision policy toward risk sensitivity rather than expected value maximization.

Keywords: rare events; reward; risk; decision making; reinforcement learning

Introduction

High rewards are often accompanied by high risks, as exemplified by venture capital investments in startups or the innovation of cutting-edge technologies. Previous research suggests that, when humans experience rewards sequentially, they tend to avoid options returning large and rare rewards. This tendency has been attributed to underweighting the impact of rare rewards (Barron & Yechiam, 2009; Camilleri & Newell, 2011; Garcia et al., 2021; Hertwig et al., 2004; Szollosi et al., 2019). However, the underweighting perspective is confounded with an explicit risk avoidance tendency; people tend to avoid options whose rewards have greater variance, which is a feature for options with large rare rewards (Juechems et al., 2017).

Here, to compare the underweighting and risk-aversion accounts, we implemented a two-armed bandit paradigm where one of the options had a higher expected value. The two options delivered either "large and rare" (LR) rewards or "frequent and medium" (FM) rewards. We found that unlike FM options, choice rates for the LR options were below chance level and were not influenced by their expected value. Computational modeling revealed that these infrequent (i.e. below chance) choices for LR options were driven by a fixed, value-independent sampling probability rather than the (learned) expected value difference between the options. We conclude that the presence of large and rare rewards disrupts value-based decision-making by making relative risk the main determinant of choice.

Results

Thirty-two participants took part in the two-armed bandit task via Prolific (https://www.prolific.com/). On each trial they were asked to choose between two stimuli to maximize their cumulative reward over 50 trials in each block. Following each choice, the rewards of both the chosen and the unchosen option were shown (Fig. 1A). Participants performed 14 blocks in total, with each block having different reward distribution pairings.

The expected value (EV) of the "better" option (i.e., higher-EV option) was 36, while that of the "worse" was 27 (i.e., lower-EV option). The "large and rare" rewards were drawn from a binomial distribution delivering large rewards (i.e., 360 when it was the better and 270 when it was the worse) in 10% of trials, and zero reward for the rest of trials. Rewards

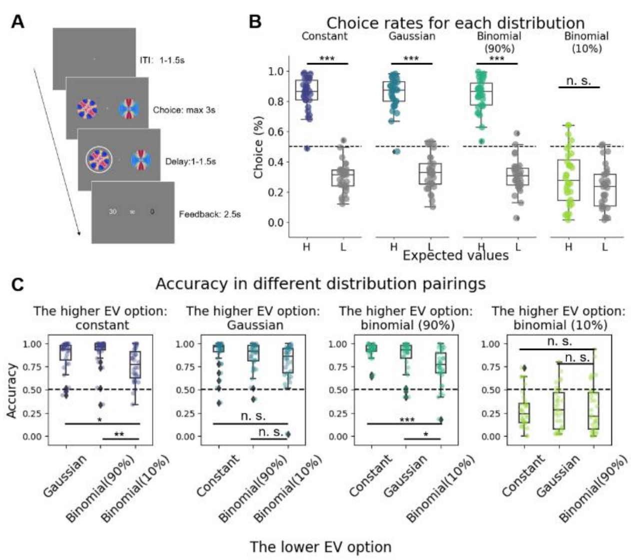

Figure 1. Task and behavioral results. (**A**) Trial structure. (**B**) Mean choice rates averaged across conditions of the higher (H) and lower (L) expected value for each distribution. (**C**) Accuracy of choosing the higher EV option in different distribution pairings.

for "frequent and medium" options were drawn from one of the following distributions: a binomial distribution (90% reward probability), a Gaussian distribution, or a constant reward structure. Participants also completed Gaussian-Gaussian pairings with unequal standard deviations, but these data were excluded from the analysis. The analysis included six distribution pairs, each appearing once as the higher-EV option, resulting in 12 conditions (Fig. 1C).

We tested whether the expected values and reward distributions affected choice rates. The overall choice rates of binomial (10%) options were significantly lower than that of all other three distributions (F(3, 93) = 68.564, p < 0.001, partial η^2 = 0.689) (Fig. 1B). Although choice rates for FM options were sensitive to EV, the choice rates of binomial (10%) options were not influenced by their expected value (distribution * EV: F(3, 93) = 204.129, p < 0.001, partial η^2 = 0.954), such that the choice rates for the high-EV and low-EV binomial (10%) options were nearly identical (pos hoc t-test with Bonf-corrected p-value: t(31) = 2.308 p = 0.111, Cohen's d = 0.389) (Fig. 1B). These results suggest general avoidance and EV insensitivity for options with large but rare rewards. The above results could be consistent with underweighting rare events.

However, if the behavioral patterns were solely driven by underweighting, one would expect higher accuracy in choosing the higher-EV option when it is paired with a lower-EV binomial (10%) option as opposed to when it is paired with a lower-EV FM option. This is because a lower-EV binomial (10%) option should be undervalued relative to FM options. Surprisingly, however, we found the lower-EV binomial (10%) option was more disruptive to accuracy than lower-EV FM options (Fig. 1C). In conditions where the binomial (10%) options had the lower EV, participants had significantly lower accuracies relative to conditions where the lower EV options delivered FM rewards (Constant: F(2) = 9.096, p < 0.001, partial η^2 = 0.227; Gaussian: F(2) = 3.262, p = 0.045, partial η^2 = 0.095; binomial (90%): F(2) = 12.837, p < 0.001, partial η^2 = 0.293) (Fig. 1C). These findings challenge a purely value-based account, which attributes lower choice rates of LR options to underweighting rare rewards

To better understand what drives people's choices in the presence of LR options, we fitted eight reinforcement learning models comparing four learning processes and two decision processes (Fig. 2A). The model family with the risk weight (w, capturing the balance between value-based and value-independent influences on decision-making) and the risk tendency for binomial (10%) options (r, reflecting the signed strength of value-independent bias, with positive values indicating more likely to choose binomial (10%) options) outperformed other models (Exceedance probability > 0.999) (Fig. 2B). The best fitting risk weight for binomial (10%) was significantly above 0.5 (t(31) = 3.046, p = 0.005, Cohen's d = 0.538) (Fig. 2C). These modelling results suggest that participants shift their decision strategy from relying on the evaluation of expected value to a heuristic risk sampling bias.

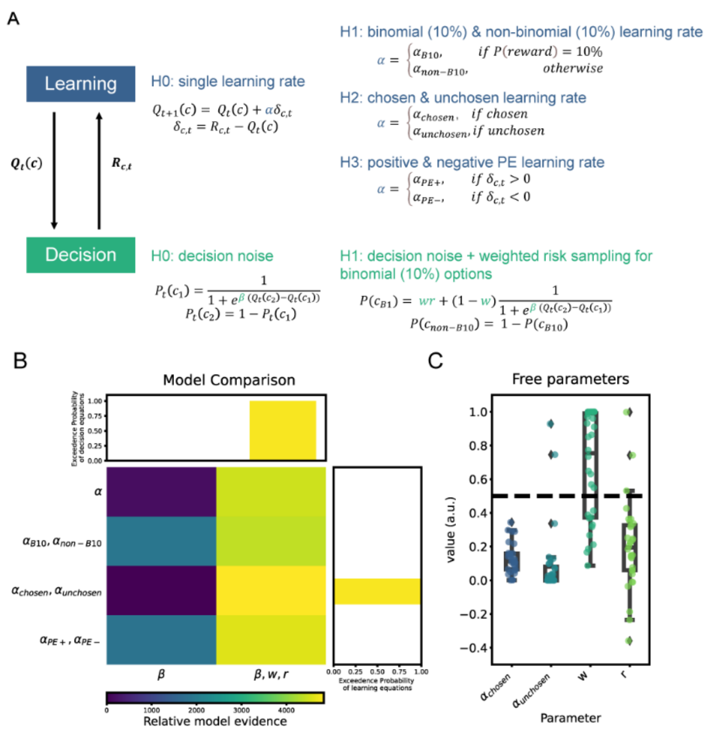

Figure 2. Modeling results. (**A**) The model space consists of eight models with distinct learning and decision processes. (**B**) Family-level model comparison. (**C**) The best-fitting free parameters.

Acknowledgements

This work was supported by scholarship 8103 EPSRC DTP22-24, awarded to LH. ANR was supported by an Australian Research Council Future Fellowship (FT230100119). YC was supported by Marie Skłodowska-Curie Fellowship (101154160 DnReLu). KT was partly supported by the EU Horizon 2020 Research and Innovation Program (ERC starting grant no. 802905).

References

Barron, G., & Yechiam, E. (2009). The coexistence of overestimation and underweighting of rare events and the contingent recency effect. *Judgment and Decision Making*, *4*(6), 447–460. https://doi.org/10.1017/s1930297500003983

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin and Review*, *18*(2), 377–384. https://doi.org/10.3758/s13423-010-0040-2

Garcia, B., Cerrotti, F., & Palminteri, S. (2021). The description-experience gap: A challenge for the neuroeconomics of decision-making under uncertainty. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1819). https://doi.org/10.1098/rstb.2019.0665

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–539. https://doi.org/10.1111/j.0956-7976.2004.00715.x

Juechems, K., Balaguer, J., Ruz, M., & Summerfield, C. (2017). Ventromedial Prefrontal Cortex Encodes a Latent Estimate of Cumulative Reward. *Neuron*, *93*(3), 705-714.e4. https://doi.org/10.1016/j.neuron.2016.12.038

Szollosi, A., Liang, G., Konstantinidis, E., Donkin, C., & Newell, B. R. (2019). Simultaneous Underweighting and Overestimation of Rare Events: Unpacking a Paradox. *Journal of Experimental Psychology: General*, 1–32. https://doi.org/10.1037/xge0000603