Title
Framing the Exploration-Exploitation Trade-Off: Distinguishing Between Minimizing Losses and Maximizing Gains

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Danwitz, Ludwig
Hosch, Ann-Katrin
von Helversen, Bettina

Publication Date
2024

Peer reviewed

Framing the Exploration-Exploitation Trade-Off: Distinguishing Between Minimizing Losses and Maximizing Gains

Ludwig Danwitz (danwitz@uni-bremen.de)

Ann-Katrin Hosch (ann-katrin.hosch@uni-bremen.de)

Bettina von Helversen (b.helversen@uni-bremen.de)
Institute for Psychology, Hochschulring 18
Bremen, 28359 Germany

Abstract

To successfully minimize losses or maximize gains, individuals must acquire a profound understanding of the rules and regularities in their environment. The current project centers on the impact of the environment on exploration and exploitation behavior. Therein, we compare costly exploration in environments, in which it is only possible to win (even though the size of the gains differs), only possible to lose, and mixed environments, in which one can win and lose. Participants engaged in a Multi-Armed Bandit task in three such conditions. Notably, participants exhibited reduced exploration in the gain domain compared to the loss domain, with the mixed domain falling in between. Interestingly, participants performed best in the mixed domain. Computational modeling of participants' choice behavior revealed that individuals tended to underestimate outcomes of unchosen options in the gain domain and overestimated them in the loss domain. We discuss two explanations for this pattern of findings: Either, effects are driven by the absolute difference between gains and losses or by the relative difference that individuals experience in relatively better or worse environments compared to their expectations (e.g., compared to previous blocks).

Keywords: exploration-exploitation trade-off; loss aversion; computational modelling; prospect theory, decisions from experience

Introduction

Consider a scenario where a group of explorers ventures into uncharted territory. How do they decide when to set up a base camp, and when to explore different regions in order to find valuable resources? These decisions are likely influenced by the decision environment. Some regions may foster abundance, with all discovered resources proving valuable, while others may harbor harmful elements such as poisonous plants or contaminated water, posing health risks. In these environments, the focus shifts from maximizing gains to minimizing losses.

Similar decision-making dynamics unfold in our everyday lives, from trying new restaurants in our favorite vacation area to investing in the stock market, where we continually decide when to explore novel options and when to exploit known ones. The balance required in these situations is commonly referred to as the exploration-exploitation trade-off and is frequently studied through Multi-Armed Bandit (MAB) tasks (Speekenbrink, 2022). In these tasks, participants have a set number of trials to choose from several options, each associated with its own reward distribution. The observed exploration and exploitation behavior in these tasks is influenced by individual and social factors, along with various environmental characteristics (cf. Mehlhorn et al., 2015). Despite the prevalence of MAB task studies, many focus exclusively on the gain domain, where the objective is to maximize reward (Chakroun et al., 2020; Daw et al., 2006; Wiehler et al., 2021, but see also Speekenbrink & Konstantinidis, 2015). However, real-world situations often involve the goal of minimizing losses, as highlighted in our introductory example. Prospect Theory (Kahneman & Tversky, 1979) underscores that behavior can significantly differ in the domain of losses, with individuals generally exhibiting loss aversion – a heightened sensitivity towards losses compared to gains. Specifically, individuals assess outcomes by comparing them to a reference point, determining whether the result is a "loss" (i.e., relatively lower than the reference point) or a "gain" (i.e., relatively higher than the reference point). These outcomes are incorporated into a value function, which exhibits greater steepness for losses than gains, signifying that losses carry more weight than gains.

The occurrence of loss aversion has been demonstrated in numerous studies across disciplines (A. L. Brown et al., 2021). In their original prospect theory, Kahneman and Tversky (1979) postulate that the experience of losing money is related to greater aggravation than the pleasure of gaining the same amount, due to the severity of the changes in the everyday life associated with either scenario. Hitherto, the question of which processes underly loss aversion is not fully answered: Some research emphasizes the role of affective intensity in shaping loss aversion (McGraw et al., 2010; Mukherjee et al., 2017). Alternatively, Yechiam and Hochman (2013) suggest that loss aversion may stem from heightened attention to losses compared to gains. Research on loss aversion predominantly utilizes lottery tasks with explicit descriptions of probabilities and outcomes (for a Meta-analysis, see A. L. Brown et al., 2021). In these "decision from description" (DfD) tasks, loss aversion consistently manifests (A. L. Brown et al., 2021, but see also Walasek & Stewart, 2021). However, effects observed in DfD tasks often do not generalize to "decision from experience" (DfE) tasks, where information is not explicitly

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

provided but must be inferred through sampling options during an exploration phase (Hertwig et al., 2004).

To derive hypotheses on the impact of the gain versus loss domain on behavior in bandit tasks, we do two steps: First, we simulate behavior under the assumption that agents show loss averse behavior in line with the formalization of prospect theory. Second, we offer a concise overview of the existing literature on the exploration of environments involving gains and/or losses including DfE tasks.

We simulated the behavior of an agent exhibiting loss aversion in a MAB task. The agent had the option to choose from ten different alternatives ("arms"), each characterized by distinct underlying reward distributions: In the mixed condition, the means of the options were randomly sampled from a normal distribution with a grand mean of 0 points and a standard deviation of 10 points. When an option is selected, a number is drawn from a normal distribution with this option's mean and a standard deviation of 10 points. The grand mean was 50 in the gain domain and -50 in the loss domain. We opted to employ a widely used computational model for MAB tasks—the Kalman SM model (Daw et al., 2006) – as the agent in this simulation. Within this model, we formalized the proposed loss aversion by transforming experienced rewards according to prospect theory where parameter λ_s reflects the agent's loss aversion. This model evaluates the agent's (random) exploration behavior using the parameter β_s. This parameter characterizes exploration actions that deviate from selecting the option with the highest calculated value.

Specifically, in the first trial, the agent starts with an initial expectation (μ_1) and an initial uncertainty regarding this expectation (σ_1^2) for any option a. The mean expected value $\mu_{c_t,t}$ and the uncertainty $\sigma_{c_t,t}$ of the chosen option c are updated on each trial t based on the prediction error δ_t and the Kalman gain K_t (1a). The Kalman gain (1c) depends on the observation variance σ_o^2 (1b) and increases with the option's uncertainty.

$$\hat{\mu}_{a=c_t,t}^{post} = \hat{\mu}_{a=c_t,t}^{pre} + K_t \cdot \delta_t \text{ with } \delta_t = r_t - \hat{\mu}_{a=c_t,t}^{pre} \quad (1a)$$

$$\hat{\sigma}_{a=c_t,t}^{2post} = (1 - K_t) \cdot \hat{\sigma}_{a=c_t,t}^{2pre} \quad (1b)$$

$$K_t = \hat{\sigma}_{a=c_t,t}^{2pre} / (\hat{\sigma}_{a=c_t,t}^{2pre} + \sigma_o^2) \quad (1c)$$

The decision is made based on the probability $P_{a,t}$ for choosing an option a on trial t, which follows a softmax function (2):

$$P_{a,t} = softmax\left(\beta_s \cdot (\lambda_{a,t} * \hat{\mu}_{a,t}^{pre})\right) \quad (2)$$

$$\lambda_{a,t} = \begin{cases} \lambda_s, if\ \hat{\mu}_{a,t}^{pre} < 0 \\ 1 \quad, if\ \hat{\mu}_{a,t}^{pre} \geq 0 \end{cases} \quad (3)$$

We fixed σ_1^2 and σ_o^2 to their true values (10^2) and μ_1 to the environments' grand means. We simulated a total of 13500 agents, of which 300 each shared the same specification of λ_s (0.5,1,1.5,2,2.5) and β_s (0.2,0.3,1) and did the task in one of the three domains. Unsurprisingly, the simulation reveals that, for all values of β_s, loss aversion (3) has no impact in the gain domain. However, loss-averse agents in the loss domain tend to switch less, i.e. exhibit less exploration behavior. Interestingly, there is only minimal impact of loss aversion in the mixed domain. Depending on their general exploration tendency β_s, and the strength of their loss aversion λ_s, this change in exploration behavior may either enhance or diminish performance: highly explorative agents benefit from loss aversion, preventing excessive exploration, while loss aversion is detrimental for highly exploitative agents, leading to even more exploitation in the loss domain (see Figure 1). In sum, the simulation indicates that a loss averse agent would be less explorative in an environment where only losses occur (loss domain) compared to an environment with only gains (gain domain) or a mixed domain where both gains and losses can occur. [1]

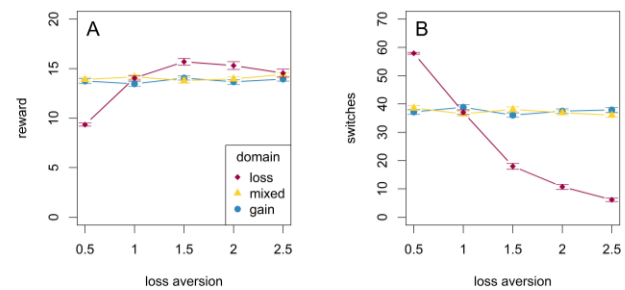

Figure 1: Simulation of Impact of Loss Aversion on Exploration Behavior and Performance in MAB tasks

Note. Rewards are adjusted relative to the grand mean of each domain. The grand mean served as the expectation on unchosen options, β_s = 0.3.

Turning to existing literature on loss aversion, the impact of loss aversion in studies involving exploration may depend significantly on whether exploration is cost-free or costly (e.g., Lejarraga & Hertwig, 2017). In cost-free exploration, individuals first undergo a sampling phase where they can explore an environment and sample from different options, and the outcomes of their samples do not influence their overall outcome. While there may be a natural inclination to cease sampling at some point, there is no disadvantage to sampling negative options during this phase (e.g., Lejarraga et al., 2012). Information search and usage are distinctly separated, and no trade-off between the two occurs. In cost-free exploration, the motivation to avoid losses should result in prolonged exploration behavior, aligning with empirical findings in such scenarios (Gurney et al., 2023; Lejarraga et al., 2012).

In contrast, scenarios involving costly exploration, such as partial feedback designs (e.g., Lejarraga & Hertwig, 2017)

[1] The simulation, can be replicated using the provided code accessible at OSF: https://t.ly/OSXUM

and MABs (e.g., Daw et al., 2006), introduce a different dynamic. In these tasks, each sample directly impacts the individual's outcome. Thus, individuals must consider avoiding excessive exploration and unfavorable options, leading to the exploration-exploitation trade-off (Cohen et al., 2007; Mehlhorn et al., 2015). In tasks with costly exploration, the picture of how different domains influence exploration becomes less clear. For instance, Chin et al. (2023) investigated an environment where participants decided whether to exploit a familiar location, retreat to a previously chosen adjacent spot, or explore a yet unchosen adjacent location, conducting the experiment in a gain and in a mixed domain. Chin et al. (2023) found that participants exhibited both decreased exploration when attempting to avoid potential losses and increased exploration when experiencing losses. Supporting the hypothesis that people explore more in the loss domain, in a DfE task with partial feedback, Lejarraga and Hertwig (2017) found that participants switched more often between risky and safe options when confronted with losses compared to the gain domain.

There are also studies employing MAB tasks, characterized by costly exploration, which reveal relatively minor differences in exploration behavior between gain and loss domains. Aberg et al. (2022) observed similar exploration behavior in participants across gain and loss conditions in an fMRI study focused on trait anxiety. Krueger et al. (2017) found that individuals engage in both directed and random exploration, irrespective of whether they are maximizing gains or minimizing losses, demonstrating agreement in exploration parameters across domains. However, their results also uncovered an overall bias towards the more uncertain option in the loss domain.

In our current study, our objective is to investigate how in a task with costly exploration individuals navigate the trade-off between exploration and exploitation when maximizing absolute gains compared to minimizing absolute losses. Additionally, we investigate whether the behavior in the mixed domain deviates from behavior in either of the other domains. Considering our simulation, one might assume that participants will exhibit the most exploration in the gain domain, followed by the mixed domain, and the least in the loss domain. This suggests a potentially myopic approach to exploration, emphasizing the numerous early losses while overlooking the potential for information gathering during exploration that could, in turn, prevent future losses.

However, our simulation is based on a formalization of loss aversion according to prospect theory, and the predictions of prospect theory often do not generalize to DfE tasks and possibly not to other tasks involving exploration such as the MAB task. Moreover, empirical evidence on exploration behavior in the loss domain varies across studies.

Given these conflicting indications from theory and empirical findings, we consider differences between domains in either of the expected directions as valuable for advancing theorizing on exploration and loss aversion and for disentangling the processes underlying each.

To further investigate how loss aversion leads to potential differences in exploration behavior between domains, we assessed participants' individual loss aversion using a description-based lottery task. Given relatively stable individual differences in loss aversion (Glöckner & Pachur, 2012), we anticipate that individuals behaving loss-averse in one task will exhibit a similar loss-averse tendency in another task. We hypothesize that the disparity in exploration behavior between domains, particularly the differences in switches between the loss and gain domains, correlates with variations in individual loss aversion. Accordingly, we will test for correlations between loss aversion and the modeled exploration tendency, as well as evaluate participants' performance.

Methods

Sample

Before data collection, we preregistered our study on the Open Science Framework (OSF).[2] We then collected data from 70 participants online on the platform prolific.co, ensuring informed consent prior to the commencement of the experiment. Our sample was limited to UK based participants to guarantee English language proficiency and maintain comparability in the potential earnings from the bonus payment for the subjects. On average, the study took approximately 21 minutes, and participants received a lump sum payment of 2.75 GBP, along with an additional bonus of on average 3.23 GBP.

The determination of the sample size was based on the analysis of a key hypothesis, specifically the differences in the number of switches between the gain, mixed, and loss domains. We assumed an effect size of f = 18 (a small to medium-sized effect) as the smallest effect of interest for this hypothesis. To detect an effect of this size in a repeated measures ANOVA with a power of beta=0.95, a sample size of 66 subjects was required.

We excluded 8 participants based on our preregistered exclusion criteria: 5 due to performance below chance level in the MAB task and 3 participants self-reported potential data flaws.

Design and Experimental Task

We implemented a MAB experiment with a single within-subject factor (domain), manipulating it as different distributions of mean reward. We used the same reward distributions as in the simulation.

Participants first completed 15 trials to familiarize themselves with the experimental task. Subsequently, they were informed that the experiment commenced and completed an initial practice block of 70 trials in the mixed domain, excluded from later analyses. Previous research indicates that participants' exploration behavior differs

[2] https://t.ly/6Hlra

between the first and subsequent blocks, as they are still familiarizing themselves with the task (Danwitz & von Helversen, n.d.) For the main experiment, participants completed six blocks of 70 trials each in randomized order, including two blocks for each domain.

In each trial, participants viewed ten boxes on the screen and chose one to draw from, with unlimited decision time. Clicking on a box revealed the associated points underneath. The top of the screen displayed the current block's status, including the number of choices made, the total choices available in the block, and the points earned in that block.

Participants were instructed to maximize accumulated points, and informed that some boxes were generally more beneficial. They were informed that the relative qualities of the boxes changed with each new block ("reset to random states") and that absolute quality varied between blocks.

Participants started with 2,500 points, and for every additional 125 points scored above 5,000, they would receive 0.1 GBP. This reward scheme was thoroughly explained.

Following the bandit task, we measured loss aversion using a game called coin toss lottery where participants chose between two gambles, one with increasing chances of losses (adapted from Brink & Rankin, 2013). Participants chose ten times between two lotteries. Both options provide a chance of 50:50 for winning or losing, like when tossing a fair coin. Option A had the chance of gaining 2000 points or incurring initially 560 points of loss, however, this value increased for each decision up to 2800 in the tenth decision. Option B offered a consistent chance of gaining or losing 400 points. One decision was randomly selected and impacted the participants' total score. We used the number of decisions for Option B as behavioral measure of loss aversion. The experiment was programmed in PsychoPy (Peirce et al., 2019).

Results

The data analysis was conducted in R, Version 4.3.2 (R Core Team). The modeling was conducted using Stan, Version 2.34 (Stan Development Team, 2024), a program for hierarchical Bayesian modeling which uses MCMC optimization.

The reported hypotheses tests were conducted in accordance with our preregistration, while the computationally modelling was performed exploratively. The data and model code are accessible at OSF.[3]

Modeling

We employed the Kalman SM model to model participants' choices, aligning closely with the agent in our simulation. The model fitting was performed separately for each block. Notably, we did not explicitly model loss aversion (λ) due to challenges in properly disentangling λ from the determination parameter β. The initial uncertainty regarding options (σ_1) and the observational noise (σ_o) were fit to their true values (10), consistent with the approach taken in several previous studies utilizing such models (Chakroun et al., 2020).

The parameter β (determination/inverse random exploration) was freely fitted in the model. In addition, we fitted μ_1 freely in the model to check whether the expected value for the unchosen options deviated from the true value. We drew 2000 post burn-in MCMC samples and did not apply thinning. We considered the procedure successful as all chains converged, with a convergence index $\hat{R} < 1.005$.

Influence of Domain on Exploration and Performance

We examined the impact of domain on exploration behavior, as measured by the number of trials where participants chose a different option than in the preceding trial (switches), the determination parameter β, the expectation on unchosen values μ_1, and the points scored by participants (reward).

Individual within-subjects ANOVAs were fitted for each dependent variable, with the condition serving as the independent variable (factor levels: loss, mixed, or gain domain).

Table 1: Effect of domain on switches, reward, determination, and expectation regarding unchosen options

Dependent variable	$F(2,368)$	p	η^2	Contrasts (p-values)		
				Loss - Gain	*Loss-Mixed*	*Mixed-Gain*
Switches	11.48	<0.001	0.06	<0.001	<0.001	0.358
Reward	14.51	<0.001	0.07	0.008	0.007	<0.001
Determination (β)	5.06	0.007	0.03	0.002	0.022	0.446
Initial Expectation (μ_1)	28.59	<0.001	0.13	<0.001	0.10	<0.001

Note. Effects are calculated within subjects. No alpha error correction was applied.

[3]https://t.ly/OSXUM

The results are summarized in Table 1 and presented in Figure 2. Reward was adjusted relative to the grand mean of a domain (Panel A). Performance was best in the mixed domain, followed by the loss domain, with the least favorable outcomes observed in the gain domain. Participants exhibited a higher number of switches in the loss domain compared to the other domains (Panel B). Additionally, participants demonstrated less random exploration, indicated by a higher determination parameter β in the mixed and gain domains compared to the loss domain (Panel C). Interestingly, we also found an effect of the domain on the expected value for the unchosen options μ_1: Relative to the true values, participants tended to underestimate the value of unchosen options in the gain domain and overestimate it in the loss domain (Panel D).

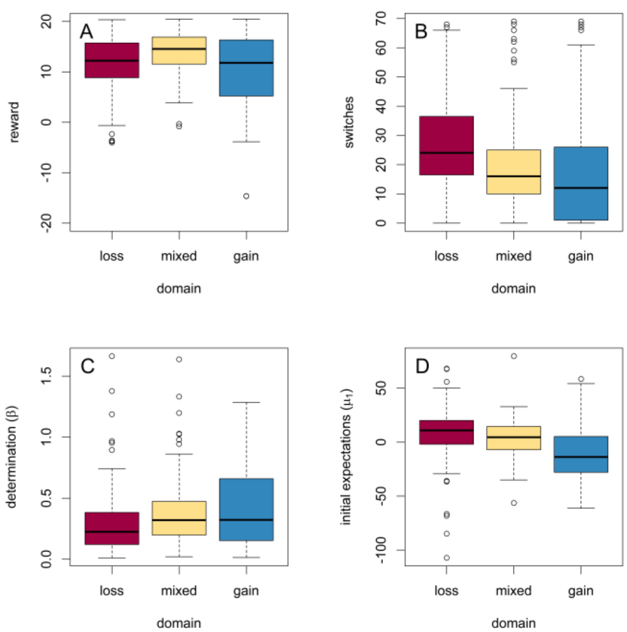

Figure 2: Effect of domain on reward, switches. determination and expectation regarding unchosen options
Note: A and B describe the impact of domain on behavioral measures; C and D depict the impact on modeled measures

Table 2: Correlations of the differences in variables between domain and behavioral loss aversion

Variable	*Loss - Gain*	*Loss - Mixed*	*Mixed - Gain*
Switches	0.31	0.36	0.00
	(p = 0.016)	(p = 0.004)	(p = 0.989)
Reward	0.28	-0.21	0.39
	(p = 0.027)	(p = 0.107)	(p = 0.002)

Note. Pearson's product-moment correlation, Number of observations in each test:62, df = 60

To assess how the behavioral loss aversion measure aligns with observed behavioral data, correlation analyses were conducted between the differences in these variables and the behavioral loss aversion measure obtained from the coin toss gambling task. Table 2 reveals significant positive correlations between behavioral loss aversion and differences in switches between the loss and the gain as well as the loss and the mixed domain. Similarly, differences in rewards between the mixed and gain as well as the loss and the gain domain correlated significantly positively with behavioral loss aversion. However, there were no correlations observed between differences in the modeled measures β and μ_1 and behavioral loss aversion.

Discussion

The current study seeks to investigate how the participants' exploration behavior varies between situations where they aim to minimize losses, situations where they strive to maximize gains, and situations in which both, gains and losses occur. Through a simulation, we demonstrated that participants, if exhibiting loss aversion according to prospect theory, would likely reduce their exploration behavior in loss environments compared to environments where they gain rewards or experience both gains and losses. However, a review of previous studies on the influence of loss aversion on exploration behavior rather suggests that losses often lead to increased exploration: While Chin et al. (2023) found that the fear of potential losses can reduce exploration, most studies indicate that people explore in the face of losses. Moreover, we observed that individuals explore more when exploration is cost-free (Gurney et al., 2023; Lejarraga et al., 2012), compared to situations where it is costly and needs to be balanced with exploitation (Krueger et al., 2017; Lejarraga & Hertwig, 2017; Yechiam et al., 2015).

In our experiment, we anticipated observing differences between domains in both, behavioral measures and modeled exploration tendencies. Additionally, we anticipated that the domain would have an impact on participants' performance. Contrary to the results of the simulation but consistent with most of the literature, participants demonstrated higher levels of exploration in the loss compared to the mixed and gain domains. This was evident in both behavioral exploration (switching between options) and the modeled measure in the MAB task.

Interestingly, participants' expectations regarding the value of unchosen options varied significantly between domains, particularly in the gain domain where participants tended to underestimate the quality of unchosen options. Notably, participants most accurately estimated the expected values of unchosen options in the mixed domain, allowing them to adapt their exploration-exploitation behavior and achieve high performance in this domain. Conversely, underestimating the values of yet unchosen options proved detrimental to exploration behavior and, consequently, performance in the gain domain.

As expected, behavioral loss aversion, measured in the coin toss gamble, correlated with differences between domains in behavioral exploration and reward in the MAB task. The results supported the notion that individuals exhibiting more loss aversion in the behavioral loss aversion measure also explored more in the MAB task.

Two plausible explanations for these results emerge: Firstly, absolute losses, in contrast to absolute gains, might drive participants to engage in restless exploration, potentially to a degree where exploration becomes detrimental to their performance. Secondly, the impact of the domain may not revolve around absolute positive or negative values but rather on the environment being relatively better or worse than expected.

The first explanation assumes, in line with prospect theory, that absolute losses are processed differently from absolute gains. Contrary to the prediction of less exploration in the loss domain, the phenomenon of loss aversion may however manifest in a way that individuals engage in more random exploration in the loss domain compared to the mixed and gain domains. The observed increased exploration behavior in cost-free exploration setups (Gurney et al., 2023; Lejarraga et al., 2012) could be explained by individuals exerting significant effort in exploration to minimize potential later losses. However, in our setup, exploration was costly, and participants needed to navigate the exploration-exploitation trade-off. Therefore, the increased exploration in the loss domain suggests that participants not only had a higher motivation to explore but also exhibited reluctance towards losses occurring during exploration. Our findings align with those by Lejarraga and Hertwig (2017), who also observed increased exploration in the loss domain in a costly DfE task. They propose that such a pattern can be explained by explaining loss aversion as an outcome of attention and arousal: if the threat of a loss induces arousal or increased attention, participants may put more effort into exploring the loss domain (Lejarraga & Hertwig, 2017; Yechiam & Hochman, 2013, 2014). However, this explanation only partially accounts for our pattern of results. In contrast to other approaches (Yechiam et al., 2015), in our case, the observed increased random exploration had a detrimental effect on participants' performance. This is difficult to explain with a better focus on the experiment's incentive structure in the loss domain, especially considering that increased random exploration is associated with noisy, non-goal-directed behavior (Wyart & Koechlin, 2016). On the other hand, our pattern of results could be explained by distress as the mediating factor: it is well-known that losses induce stress or anxiety in participants (Hayes & Wedell, 2020; Hochman & Yechiam, 2011) and stress can be detrimental to performance in exploration-exploitation scenarios (Aylward et al., 2019; V. M. Brown et al., 2022; Kool et al., 2017)

A second plausible explanation for our findings revisits the observation of biased expectations regarding the value of unchosen options. If participants find themselves in a surprisingly rich environment while maintaining a medium aspiration level, it may translate into a low aspiration level. To test whether this approach could explain our pattern of results, we repeated our initial simulation with either overly optimistic, realistic, or pessimistic expectations regarding unchosen options (Figure 3). The results of the simulation share all key features of our empirical data.

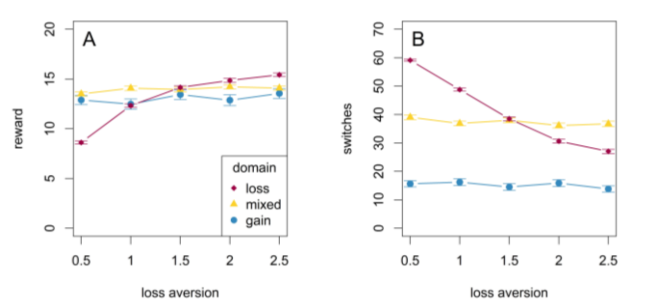

Figure 3: Simulation of Impact of Loss Aversion and Biased Initial Expectation on Exploration Behavior and Performance in MAB tasks

Note. Rewards were adjusted to points above grand mean. The grand mean served as the expectation on unchosen options, β_s = 0.3, μ_1 resembled the grand mean in the mixed domain (0), while it was by 15 points too optimistic in the loss domain (-50+15) and too pessimistic in the gain domain (50-15).

This explanation also aligns with models suggesting that exploration occurs when results fall below a certain threshold (Kohno & Takahashi, 2017; Newell & Lee, 2011). Empirical evidence also supports the idea that high state (Richner et al., 2023) and trait (Ferecatu & De Bruyn, 2022; Schwartz et al., 2002) aspiration levels lead to increased exploration. Additionally, the initial block, consistently in the mixed domain, might have worked as an anchor (Furnham & Boo, 2011), introducing an initial expectation on the unchosen values. However, these explanations alone cannot fully account for the differences in determination and random exploration between domains, as well as the correlation between behavioral loss aversion and differences in switches and rewards between conditions.

Based on our current data, we cannot definitively determine which of these processes underlies our results or whether both contribute. To address this, we plan to conduct further research for clarification.

Conclusion

Our findings suggest that individuals tend to explore more when their goal is to minimize losses compared to situations where they aim to maximize gains or operate in a mixed domain involving both gains and losses. Interestingly, the best performance is observed in the mixed domain, while individuals excessively explore in the loss domain and tend to exploit even suboptimal options in the gain domain. While we observe the influence of loss aversion on these processes, the exact mechanisms underlying these behaviors warrant further investigation.

Acknowledgments

We declare that none of the authors have competing financial or non-financial interests. We received funding from the University of Bremen, Humans-on-Mars Initiative Project SE2.

References

Aylward, J., Valton, V., Ahn, W.-Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour*, *3*(10), 1116–1123. https://doi.org/10.1038/s41562-019-0628-0

Brink, A. G., & Rankin, F. W. (2013). The Effects of Risk Preference and Loss Aversion on Individual Behavior under Bonus, Penalty, and Combined Contract Frames. *Behavioral Research in Accounting*, *25*(2), 145–170. https://doi.org/10.2308/bria-50408

Brown, A. L., Imai, T., Vieider, F. M., & Camerer, C. F. (o. J.). *Meta-Analysis of Empirical Estimates of Loss-Aversion*. SSRN.

Brown, V. M., Hallquist, M. N., Frank, M. J., & Dombrovski, A. Y. (2022). Humans adaptively resolve the explore-exploit dilemma under cognitive constraints: Evidence from a multi-armed bandit task. *Cognition*, *229*, 105233. https://doi.org/10.1016/j.cognition.2022.105233

Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F., & Peters, J. (2020). Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. *eLife*, *9*, e51260. https://doi.org/10.7554/eLife.51260

Danwitz, L., & Helversen, B. V. (2024). *Observational Learning of Exploration-Exploitation Strategies in Bandit Tasks*. https://doi.org/10.2139/ssrn.4732127

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879. https://doi.org/10.1038/nature04766

Ferecatu, A., & De Bruyn, A. (2022). Understanding Managers' Trade-Offs Between Exploration and Exploitation. *Marketing Science*, *41*(1), 139–165. https://doi.org/10.1287/mksc.2021.1304

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*(1), 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Gurney, N., Miller, J., & Pynadath, D. (2023). *The Role of Heuristics and Biases in Complex Choices* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-2472194/v1

Hayes, W. M., & Wedell, D. H. (2020). Autonomic responses to choice outcomes: Links to task performance and reinforcement-learning parameters. *Biological Psychology*, *156*, 107968. https://doi.org/10.1016/j.biopsycho.2020.107968

Hochman, G., & Yechiam, E. (2011). Loss aversion in the eye and in the heart: The autonomic nervous system's responses to losses. *Journal of Behavioral Decision Making*, *24*(2), 140–156. https://doi.org/10.1002/bdm.692

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263.

Kohno, Y., & Takahashi, T. (2017). A cognitive satisficing strategy for bandit problems. *International Journal of Parallel, Emergent and Distributed Systems*, *32*(2), 232–242. https://doi.org/10.1080/17445760.2015.1075531

Kool, W., Shenhav, A., & Botvinick, M. M. (2017). Cognitive Control as Cost-Benefit Decision Making. In T. Egner (Hrsg.), *The Wiley Handbook of Cognitive Control* (1. Aufl., S. 167–189). Wiley. https://doi.org/10.1002/9781118920497.ch10

Krueger, P. M., Wilson, R. C., & Cohen, J. D. (2017). Strategies for exploration in the domain of losses. *Judgment and Decision Making*, *12*(2), 104–117. https://doi.org/10.1017/S1930297500005659

Lejarraga, T., & Hertwig, R. (2017). How the threat of losses makes people explore more than the promise of gains. *Psychonomic Bulletin & Review*, *24*(3), 708–720. https://doi.org/10.3758/s13423-016-1158-7

Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition*, *124*(3), 334–342. https://doi.org/10.1016/j.cognition.2012.06.002

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., & Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*(3), 191–215. https://doi.org/10.1037/dec0000033

Newell, B. R., & Lee, M. D. (2011). The right tool for the job? Comparing an evidence accumulation and a naive strategy selection model of decision making. *Journal of Behavioral Decision Making*, *24*(5), 456–481. https://doi.org/10.1002/bdm.703

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Richner, J., Zagorac-Uremović, Z., & Laureiro-Martínez, D. (2023). Individual and context-evoked antecedents of exploration-exploitation performance. *Frontiers in Psychology*, *14*, 1167135. https://doi.org/10.3389/fpsyg.2023.1167135

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice.

Journal of Personality and Social Psychology, *83*(5), 1178–1197. https://doi.org/10.1037/0022-3514.83.5.1178

Speekenbrink, M. (2022). Chasing Unknown Bandits: Uncertainty Guidance in Learning and Decision Making. *Current Directions in Psychological Science*, *31*(5), 419–427. https://doi.org/10.1177/09637214221105051

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and Exploration in a Restless Bandit Problem. *Topics in Cognitive Science*, *7*(2), 351–367. https://doi.org/10.1111/tops.12145

Walasek, L., & Stewart, N. (2021). You cannot accurately estimate an individual's loss aversion using an accept–reject task. *Decision*, *8*(1), 2–15. https://doi.org/10.1037/dec0000141

Wiehler, A., Chakroun, K., & Peters, J. (2021). Attenuated Directed Exploration during Reinforcement Learning in Gambling Disorder. *The Journal of Neuroscience*, *41*(11), 2512–2522. https://doi.org/10.1523/JNEUROSCI.1607-20.2021

Wyart, V., & Koechlin, E. (2016). Choice variability and suboptimality in uncertain environments. *Current Opinion in Behavioral Sciences*, *11*, 109–115. https://doi.org/10.1016/j.cobeha.2016.07.003

Yechiam, E., & Hochman, G. (2013). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive Psychology*, *66*(2), 212–231. https://doi.org/10.1016/j.cogpsych.2012.12.001

Yechiam, E., & Hochman, G. (2014). Loss Attention in a Dual-Task Setting. *Psychological Science*, *25*(2), 494–502. https://doi.org/10.1177/0956797613510725

Yechiam, E., Retzer, M., Telpaz, A., & Hochman, G. (2015). Losses as ecological guides: Minor losses lead to maximization and not to avoidance. *Cognition*, *139*, 10–17. https://doi.org/10.1016/j.cognition.2015.03.001