How losses change the way we search: valence effects on generalization-guided exploration

Ingrid Martin[1] **Charley M. Wu**[2] **Toby Wise**[1]

[1] Centre for Neuroimaging Sciences, Kings College London
[2] Centre for Cognitive Science, Institute of Psychology, Technical University of Darmstadt

Abstract

Gaussian Process (GP) regression models capture how individuals balance exploration and exploitation by generalizing from rewarding experiences to guide search. However, it remains unclear whether people generalize similarly from losses, and how the presence of losses might alter exploration strategy. Across two pilot studies, participants completed 24 blocks (8 per valence) of a valenced spatially correlated bandit task. Reward values were either all positive, all negative, or mixed, with valence manipulated within subjects across blocks. Distances between sequential clicks were greater in negative and mixed conditions than positive and local search was less frequent. These effects replicated across studies and were not associated with differences in performance, suggesting the framing of losses alters the spatial structure of search rather than its breadth or effectiveness. Preliminary modelling extends GP-UCB models to investigate whether the observed behavioural differences are better explained by asymmetric weighting of negative versus positive outcomes, or by shifts in the prior mean over expected rewards, whereby people hold more optimistic expectations for unobserved options in loss environments.

Keywords: Gaussian-process, generalization, exploration-exploitation, loss-aversion

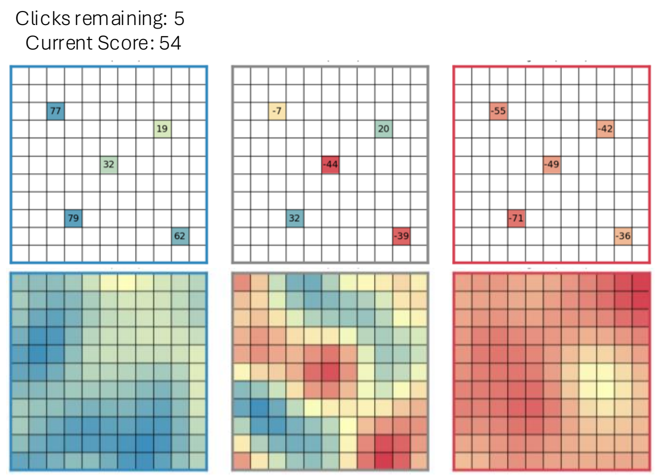

Figure 1: Valenced spatially correlated 121-armed bandit (adapted from Wu et al. 2018). Participants completed 24 rounds (8 per valence, interleaved). On each round participants had 10 clicks to reveal tiles. Conditions were signalled by border colour: blue (positive), red (negative), yellow (mixed).

Introduction

Many real-world decisions require searching through large spaces of possible actions under limited time, balancing the need to explore unfamiliar options with the desire to exploit known rewards. Gaussian process models of function learning, combined with optimistic sampling strategies such as upper confidence bound sampling (GP-UCB), provide a compelling account of how people navigate this exploration-exploitation trade-off by generalizing spatially from past experience to guide search toward promising but uncertain options (Wu et al., 2018). However, this work has largely focused on environments framed in terms of gains, leaving open the question of how people search spatially correlated spaces when outcomes involve losses — a context that is both common in real-life environments and known to alter exploration-exploitation decisions in systematic ways (Chin et al., 2023; Krueger et al., 2017). Understanding how valence shapes exploration-exploitation and generalization is particularly relevant for computational psychiatry, where disruptions to reward and loss processing are central to conditions such as depression and anxiety, and where the computational mechanisms linking outcome valence to strategic behavior remain poorly specified.

Methods

We conducted two pilot studies on healthy adults (aged 18–40) recruited via Prolific (N = 20 each). Study 2 introduced minor instruction clarifications based on participant feedback from Study 1; the task was otherwise identical. We adapted the 11×11 spatially correlated bandit task of Wu et al. (2018). Environments were

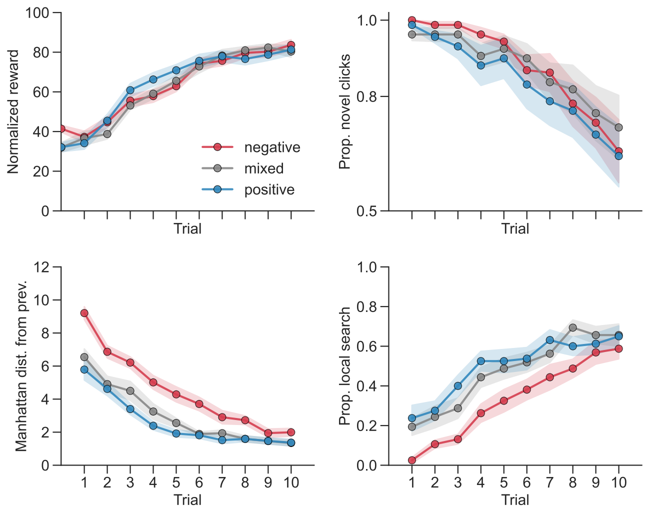

Figure 2: Behavioural measures across trials by valence condition. Participants showed matched reward across conditions, similar rates of novel exploration, but greater distances between clicks and reduced local search in negative environments. Error bands show ± 1 SEM.

generated using a Gaussian process prior (RBF kernel, length scale = 2) and scaled to three valence conditions: positive (0 to +100), negative (0 to -100), and mixed (-50 to +50). Displayed tile values included Gaussian noise ($\sigma = 1$) at each observation, so repeated clicks on the same tile could yield different values. Conditions were signalled by coloured borders (blue, red, yellow) and interleaved across 24 blocks (8 per valence). Grids were sampled without replacement and randomly assigned to conditions for each participant. On each block, participants had 10 clicks to reveal tiles on the grid, with the goal of maximising their average reward (positive), minimising average loss (negative), or both (mixed). Participants were informed that the spatial structure and value distributions were matched across conditions. Bonus payment was calculated from one randomly selected grid per valence; negative scores were subtracted from positive, preserving the aversive framing of losses. Before beginning, participants completed a comprehension quiz and practice trials to ensure understanding of the task and bonus structure. Participants were also shown examples of fully revealed grids.

Experimental Results

For brevity, results from Study 2 are reported below; all effects replicated across both datasets. Linear mixed-effects models examined the effect of valence condition (positive, negative, mixed) and trial number on behavioural measures, with random intercepts for participant. **Reward.** Normalized reward did not differ significantly between positive and negative environments ($b = 0.61$, $z = 0.53$, $p = .594$), nor did the rate of reward improvement over trials (valence $\times$ trial interaction: $b = 0.16$, $z = 0.43$, $p = .666$). Reward increased significantly across trials in both conditions ($b = 4.95$, $z = 13.11$, $p < .001$). **Novelty.** Participants in positive environments made a slightly but significantly lower proportion of novel clicks than those in negative environments ($b = -0.04$, $z = -3.55$, $p < .001$). However, there was no significant valence $\times$ trial interaction ($b = 0.001$, $z = 0.33$, $p = .743$) — novelty declined at a similar rate across conditions ($b = -0.04$, $z = -4.26$, $p < .001$). This suggests that while valence had a small effect on overall exploration of new tiles, it did not modulate how this tendency changed over time. **Search distance.** Participants in negative environments moved significantly farther between successive clicks than those in positive environments ($b = -1.90$, $z = -7.09$, $p < .001$). Both conditions showed significant distance decay over trials, but the rate of decay was steeper in negative environments (negative slope: $b = -0.75$; positive slope: $b = -0.45$; interaction: $b = 0.30$, $z = 6.05$, $p < .001$; LRT: $\chi^2(1) = 30.83$, $p < .001$). This suggests that negative outcomes promote broader spatial exploration, particularly in early trials. **Local search.** Consistent with the distance results, participants in negative environments engaged in significantly less local search (proportion of moves ≤ 1 Manhattan distance) compared to the mixed baseline ($b = -0.14$, $z = -7.81$, $p < .001$). Positive environments did not differ significantly from mixed ($b = 0.03$, $z = 1.37$, $p = .171$). Local search increased over trials across all conditions ($b = 0.06$, $z = 8.48$, $p < .001$).

Discussion

Valence framing systematically altered the spatial structure of search despite matched reward distributions and task performance. Negative environments elicited greater distances between clicks and reduced local search, while novelty differences were small. These findings suggest that losses change *how* people search rather than *how much* they explore. Preliminary modelling indicates that standard GP-UCB parameters cannot capture the observed distance differences between conditions. Future work will extend these models to incorporate loss aversion and prior mean shifts to account for these valence-driven changes in search behaviour. These samples are small pilot studies and future work will aim to replicate in larger, pre-registered samples. Importantly, disrupted exploration and loss sensitivity are central to depression and anxiety; characterising how valence shapes the computational mechanisms of search may help identify individual differences relevant to these conditions.

Acknowledgments Generative AI was used for coding and language editing in this abstract. CMW is supported by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) under Germany's Excellence Strategy (EXC 3066/1 "The Adaptive Mind"', Project No. 533717223), and the Excellence Cluster "Reasonable AI" by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) under Germany's Excellence Strategy – EXC-3057. TW is supported by a Wellcome Career Development Award, 225945/Z/22/Z.

References

Chin, A., Hagmann, D., & Loewenstein, G. (2023). Fear and promise of the unknown: How losses discourage and promote exploration. *Journal of Behavioral Decision Making*, *36*(3), e2309. https://doi.org/10.1002/bdm.2309.

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.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*, 915–924. https://doi.org/10.1038/s41562-018-0467-4.