CMML ICA1

Main Report

2074

Neurocomputational Analysis of Adaptive Strategies in Schema-Based Decision Making

Abstract

Optimizing strategies in schema-based decision making remains complex. To address this gap, we developed computational modeling to explore different adaptive strategies. Win-Stay, Lose-Shift (WSLS) improved accuracy and performance. Inhibitory performance feedback paradoxically enhanced performance via slower processing. A speed-focused 'High-Payoff' strategy with schema abandonment maximized cumulative performance, outperforming WSLS and adaptive switching strategies. It reveals crucial trade-offs between local values and overall goals.

Main

Understanding how individuals make decisions when choices are guided by learned internal representations (schemas) and influenced by task structure is crucial for cognitive science. It is influenced by factors like confidence in schema knowledge, exploration patterns, and task incentives (Speekenbrink & Konstantinidis, 2015; Gershman, 2018). Computational models provide a formal framework to investigate the cognitive mechanisms underlying performance in these complex environments. The specific paradigm simulated here involves a multi-round task where participants must select multiple items per round (Fig. 1a). A significant performance bonus is contingent upon selecting items belonging to the same underlying schema (e.g., artworks by the same artist), rewarding accurate schema identification and application during decisions. The baseline model employed simulates this schema-based decision-making over multiple rounds (Fig. 1b). It updates schema-specific and generic confidence via exploration and feedback. Decision-making is captured by a sequential evidence accumulation process, akin to drift-diffusion models (Ratcliff & McKoon, 2008). Attention shifts between items based on a weighted integration of confidence estimates and potential bonus. It supports evidence accumulation, compared against dynamic thresholds to trigger schema identification and item selection. This framework captures the interplay between schema learning and decision making.

While the baseline model simulates core learning and decision processes, human behaviour exhibits adaptive strategies during multi-round tasks not explicitly included. Strategy selection dynamics—the mechanisms by which individuals adjust and adapt task strategies—are central to this process, where previous outcome-based feedback can sharpen these strategies, with emotional states and task demands potentially altering their efficacy (Schulz et al., 2019; Lieder & Griffiths, 2020). This project aimed to investigate the impact of incorporating several adaptive strategies into the baseline computational model via simulation. By comparing these different adaptive strategies, we sought to understand their relative effectiveness in maximizing overall schema-based learning task performance.

Win-Stay, Lose-Shift (WSLS), rooted in behavioral ecology, promotes persistence with successful decisions and switching after failures, optimizing performance in iterative tasks (Fawcett et al., 2014). Building upon the baseline model, we introduced the WSLS heuristic to investigate how reinforcement history influences schema selection. This WSLS modification biases attention towards items of a specific schema in the subsequent round if it led to a "successful match" in the previous round, and biases attention away if it led to a " failed match". Simulation results indicate that incorporating the WSLS heuristic increases the frequency of consecutively selecting the same schema (Fig. 2a). WSLS interacted with decision dynamics, increasing Reaction Time (RT) / Observation counts (OB) / Attention Shifts (AS) for repeated 'successful' schemas but decreasing them for repeated 'failed' schemas (Supplementary Fig. 1). This strategy improved schema accuracy of later rounds and achieved slightly higher total performance (Fig. 2b and Fig. 2c).

The observation that the WSLS heuristic, while enhancing accuracy, also modulated decision dynamics (RT/OB/AS) hinted at potential impacts on the underlying evidence accumulation process itself. Recognizing that in realistic multi-round tasks, prior performance often influences emotional states and subsequent cognitive processing, we further extended the model to explore this. We introduced an emotion factor that directly adjusts the evidence accumulation rate based on the total performance of the preceding rounds. Simulating four states (Baseline, Feedback, Excited, Depressed) with varying incentive/inhibition strengths, we found paradoxically that only the 'Depressed' state yielded significantly higher performance compared to baseline, particularly later in the task (Fig. 2d). This improvement occurred despite markedly increased RTs for schema identification under inhibitory conditions, suggesting a benefit from slower processing (Fig. 2e).

However, persistent failure might elicit alternative adaptive strategies beyond cautious processing. If agents consistently fail to achieve schema-matching bonuses, they might perceive this goal as too difficult or time-consuming and strategically abandon it. We implemented a ' High-Payoff' strategy, which bypassed schema selection, prioritized attention to high-payoff items, and used lower decision thresholds. Simulations showed this adaptive strategy significantly boosted per-round performance, drastically reduced round duration allowing more rounds to be completed, and resulted in substantially higher total cumulative performance (Fig. 2c). Interestingly, it didn't impair schema-matching accuracy when fewer schemas were present (Supplementary Fig. 2).

To find the optimal strategy, we systematically compared WSLS, ‘High-Payoff' strategy, and the adaptive 'Combined' strategy (initial WSLS switching to ‘High-Payoff’). While all strategies outperformed the baseline, the ‘High-Payoff’ strategy consistently achieved the highest total performance across both longer and shorter simulated task durations (Fig. 2c and Fig. 2f). This was driven by both superior per-round performance and the completion of more rounds due to its inherent speed.

Our simulations demonstrate that incorporating psychologically plausible heuristics and adaptive mechanisms significantly alters decision outcomes and performance maximization strategies compared to the baseline model. The WSLS heuristic, while marginally beneficial, interacts complexly with decision dynamics, where key drivers may exist but be constrained. The counter-intuitive performance enhancement under simulated 'Depression' highlights potentially adaptive speed-accuracy trade-offs. The inhibitory feedback slowed decisions, potentially preventing premature commitments and costly errors, echoing ideas about cognitive control adjustments after negative outcomes (Shenhav et al., 2017; Cavanagh & Frank, 2019). While this increases processing time, the inhibition mechanism might functionally balance evidence accumulation, preventing choices based solely on transient value signals.

The pronounced effectiveness of the 'High-Payoff' strategy, particularly when employed proactively, underscores strategic adaptation to task structure and perceived difficulty Kool & (Wilson et al., 2018). The dominance of the Direct High-Payoff strategy demonstrates that maximizing throughput, primarily by eliminating the time-consuming schema identification stage, can be highly effective when there are few schemas. This suggests exploring a combined approach: strategically abandoning global schema matching but retaining cautious, inhibition-modulated item evaluation focused on direct payoffs. It is potentially applicable to in real-world 'Exploit-Best-Available' scenarios when the overall picture is not clear.

These findings highlight the utility of computational models for exploring cognitive strategy, revealing the place of local value evaluation in global tasks. However, a fundamental tension exists between the cautious inhibition potentially promoted by negative feedback and the aggressive schema abandonment of the High-Payoff strategy. Future work includes investigating the balance between cautious evaluation and goal simplification. To identify strategy use in the real world, fitting these models with human data is also needed. It helps to examine potential neural correlations and explore finer-grained triggers for shifts.

Figures

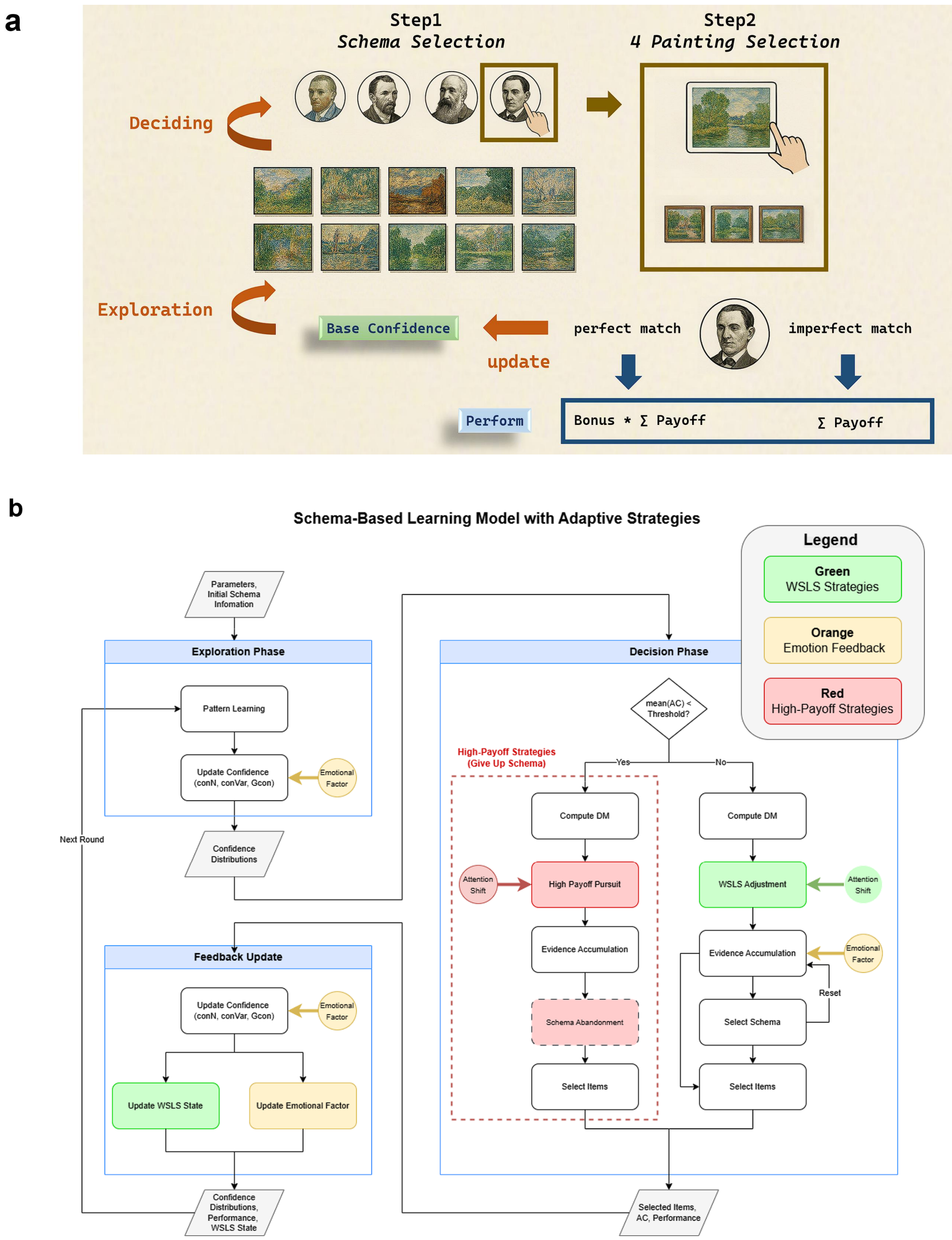

Fig. 1: The simulated tasks and the model architecture.

a, Schematic representation of the multi-round, schema-based decision-making task. **b,** Architecture of the neurocomputational model integrating different adaptive strategies. The three extension modules of WSLS, emotional feedback, and 'High-Payoff' strategy are marked in green, orange, and red, respectively.

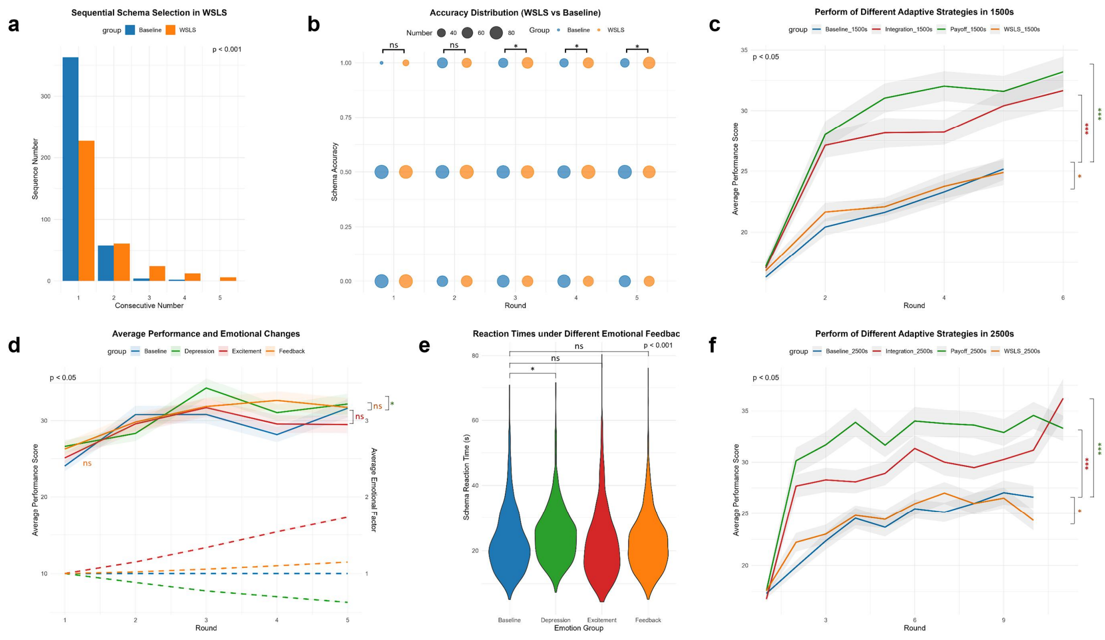

Fig. 2: Pattern learning under different Adaptive Strategies.

a, Histogram of consecutive schema selections for Baseline and WSLS (Kolmogorov-Smirnov test). **b,** Schema-matching accuracy for Baseline and WSLS across rounds. **c,** Performance of adaptive strategies (Baseline, WSLS, Integration, Payoff, Combined) over 1500 s. **d,** Performance under emotional states (Baseline, Feedback, Excitement, Depression) across rounds. **e,** Reaction times for schema identification under emotional states.

f, Performance of adaptive strategies over 2500 s. (b–f: one-tailed Wilcoxon signed-rank test vs. Baseline, Bonferroni-corrected $P < 0.0167$; see main text for details). Significance markers: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

References

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. Trends in cognitive sciences, 18(8), 414–421.

Fawcett, T. W., Hamblin, S. and Giraldeau, L.-A. (2013). Exposing the behavioral gambit: the evolution of learning and decision-making. Behavioral Ecology, 24(1), pp. 2-11.

Gershman, S. J. (2018). Deconstructing the human algorithms for exploration. Cognition, 173, pp. 34-42.

Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. The Behavioral and brain sciences, 43, e1.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural computation, 20(4), 873–922.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for Rewards Like a Child Means Less Generalization and More Directed Exploration. Psychological science, 30(11), 1561–1572.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron, 79(2), 217–240.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. Topics in cognitive science, 7(2), 351–367.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore-exploit dilemma. Journal of experimental psychology. General, 143(6), 2074–2081.