Supplementary Materials

Mini-project 2 - 2074

Supplementary Methods

Code Availability

Code are available at `https://github.com/guobiao-ye/Schema-Based_Decision_Making.git`.

Overview of the Baseline Model

The baseline model from CMML3 simulates schema-based learning and decision-making. It includes an exploration phase, where agents update schema-specific and generic confidence, and a decision phase, where evidence is accumulated to select items based on attention-driven probabilities.

Modifications in the Extended Model

WSLS Strategy

The WSLS strategy biases item selection in Phase 1 of round $r > 1$ based on the outcome of Phase 2 in round $r-1$. Two state variables are introduced: previous_decision_outcome and previous_dominant_schemaID. The outcome is:

$$\text{previous\_decision\_outcome} = \begin{cases} \text{“Win”}, & \text{if Outputs\_cho.AC}[2r-2] = 1, \\ \text{“Lose”}, & \text{otherwise.} \end{cases}$$

The dominant schema is:

$$\text{previous\_dominant\_schemaID} = \begin{cases} \text{unique(Schema\_res)}, & \text{if AC} = 1, \\ \arg\max_{\text{ID}} \text{count(Schema\_res[ID])}, & \text{if AC} < 1 \text{ and unique,} \\ \text{NA}, & \text{otherwise,} \end{cases}$$

where Schema_res lists schema IDs chosen in Phase 2. In Phase 1, attention probability for item i is:

$$p_i \propto \exp(\Phi \times \text{Item\_EI.evidence}_i \times \text{Item\_EI.DM}_i \times \text{Item\_EI.timevar}_i),$$

and adjusted by WSLS:

$$p_i \leftarrow \begin{cases} p_i \times \text{WSLS\_boost\_factor}, & \text{if outcome} = \text{“Win” and Item\_EI.Schema}_i = \text{previous\_dominant\_schemaID}, \\ p_i \times \text{WSLS\_penalty\_factor}, & \text{if outcome} = \text{“Lose” and Item\_EI.Schema}_i = \text{previous\_dominant\_schemaID}, \\ p_i, & \text{otherwise,} \end{cases}$$

with normalization $\sum_i p_i = 1$. Defaults are WSLS_boost_factor = 100 and WSLS_penalty_factor = 0.01 for WSLS, and 0 for Baseline.

Emotion-Driven Evidence Accumulation

This mechanism modulates evidence accumulation based on prior performance. For round $r > 1$, the previous round's performance is:

$$\text{last\_performance}_{r-1} = \sum_{p \in \{2r-2, 2r-1\}} \text{Outputs\_cho.performance}[p],$$

normalized by the maximum performance:

$$\text{max\_possible\_perf} = \max(\text{schemainfo.payoff}) \times 3 \times 4 \times 2,$$

$$\text{perf\_ratio}_{r-1} = \frac{\text{last\_performance}_{r-1}}{\text{max\_possible\_perf}}.$$

The emotion factor, initialized as $\text{emotion\_factor}_1 = 1$, updates as:

$$\text{emotion\_factor}_r = \begin{cases} \text{emotion\_factor}_{r-1} + \gamma_{\text{incentive}} \times (\text{perf\_ratio}_{r-1} - \theta_{\text{con}}), & \text{if perf\_ratio}_{r-1} > \theta_{\text{con}}, \\ \text{emotion\_factor}_{r-1} - \gamma_{\text{inhibition}} \times (\theta_{\text{con}} - \text{perf\_ratio}_{r-1}), & \text{otherwise}, \end{cases}$$

where $\theta_{\text{con}} = 0.3$, and $\gamma_{\text{incentive}}$, $\gamma_{\text{inhibition}}$ control feedback strength. Evidence for item i at time t is scaled:

$$\text{Item\_EI.evidence}_i(t) \leftarrow \text{Item\_EI.evidence}_i(t) + \text{emotion\_factor}_r \times \mathcal{N}(\text{Item\_EI.N}_i, \text{Item\_EI.Var}_i).$$

Conditions tested were: Baseline ($\gamma_{\text{incentive}} = 0$, $\gamma_{\text{inhibition}} = 0$), Feedback ($\gamma_{\text{incentive}} = 0.5$, $\gamma_{\text{inhibition}} = 0.3$), Excitement ($\gamma_{\text{incentive}} = 1.0$, $\gamma_{\text{inhibition}} = 0.1$), and Depression ($\gamma_{\text{incentive}} = 0.1$, $\gamma_{\text{inhibition}} = 1.0$).

High-Payoff Selection Mode

This mode activates when prior accuracy is low:

$$\text{give\_up\_mode}_r = \left(\frac{1}{2(r-1)} \sum_{p=1}^{2(r-1)} \text{Outputs\_cho.AC}[p] < \text{AC\_threshold} \right),$$

with AC_threshold = 0.25 (Combined), 2 (Direct High-Payoff), or 0 (Baseline, WSLS). In this mode, schema selection is skipped:

$$\text{Outputs\_cho.Schema}[p] = \text{“give\_up”}, \quad \text{Outputs\_cho.Schema\_RT/OB/AS}[p] = 0,$$

using lower thresholds (give_up_schema_thres, give_up_item_final_thres = 7.5). Attention probability includes payoff:

$$p_i \propto \exp(\Phi \times \text{Item\_EI.evidence}_i \times \text{Item\_EI.timevar}_i + \text{payoff\_attention\_weight} \times \text{Item\_EI.payoff}_i),$$

with payoff_attention_weight = 0.5. The decision metric is:

$$\text{DM}_i = [0.8w \times \text{Scon}_i + (1 - 0.8w) \times \text{Gcon}_i] \times \text{payoff}_i.$$

Simulation Protocol

Simulations ran for 100 agents per condition with 1500s or 2500s. Conditions were:

- Baseline: No modifications.
- Feedback: Moderate incentive, moderate inhibition.
- Excitement: High incentive, low inhibition.
- Depression: Low incentive, high inhibition.
- WSLS: WSLS strategy always active.
- Direct High-Payoff: High-payoff mode always active.
- Combined: WSLS first, and high-payoff mode active then.

Analytical Methods

Performance, AC, Schema RT, Schema OB, and Schema AS were analyzed using R scripts.

Performance and Accuracy

Mean performance per round was:

$$\text{mean\_performance}_r = \frac{1}{N}\sum_{s=1}^{N} \text{Outputs\_cho.performance}[s, r],$$

with standard error:

$$\text{se\_performance}_r = \frac{\text{sd}(\text{Outputs\_cho.performance}[*, r])}{\sqrt{N}}.$$

Accuracy was visualized via bubble plots, with bubble size reflecting frequency. Wilcoxon tests compared conditions (two-sided for Feedback/WSLS vs. Baseline; one-sided for others).

Reaction Time, Observation, and Attention Shifts

Kruskal-Wallis tests compared Schema RT across conditions. For WSLS, data were grouped by prior schema and AC, visualized via violin plots for Schema RT, Schema OB, and Schema AS. The test statistic was:

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(N+1).$$

Schema Selection Continuity

WSLS streak lengths were computed as:

$$\text{streak\_length}_s = \text{count}(\text{consecutive Schema}_p = \text{Schema}_{p-1}),$$

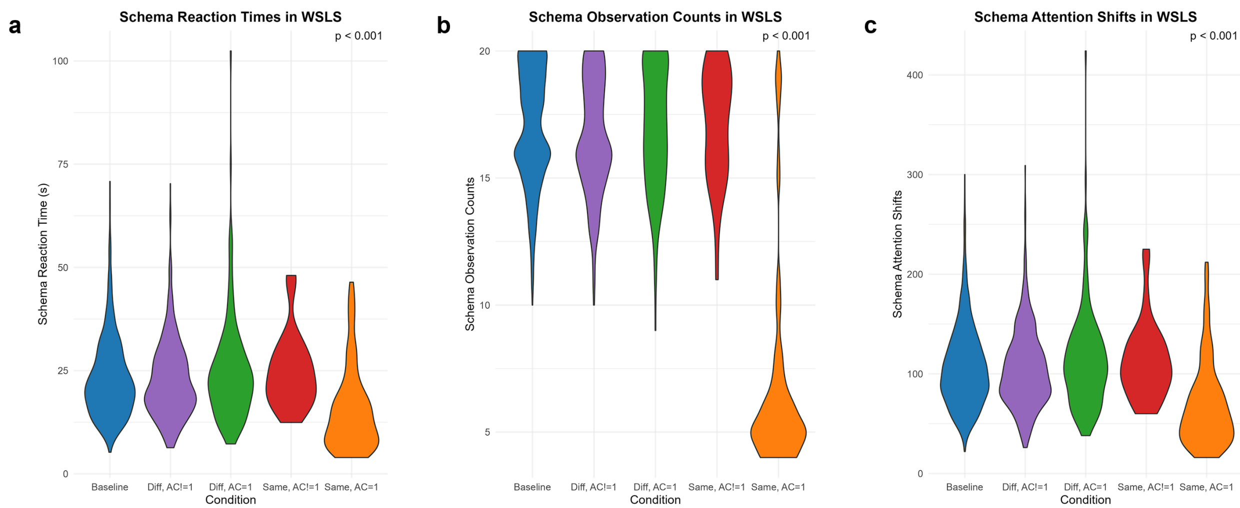
and compared to Baseline using a Kolmogorov-Smirnov test:

$$D = \sup_x |F_{\text{WSLS}}(x) - F_{\text{Baseline}}(x)|.$$

Reflection

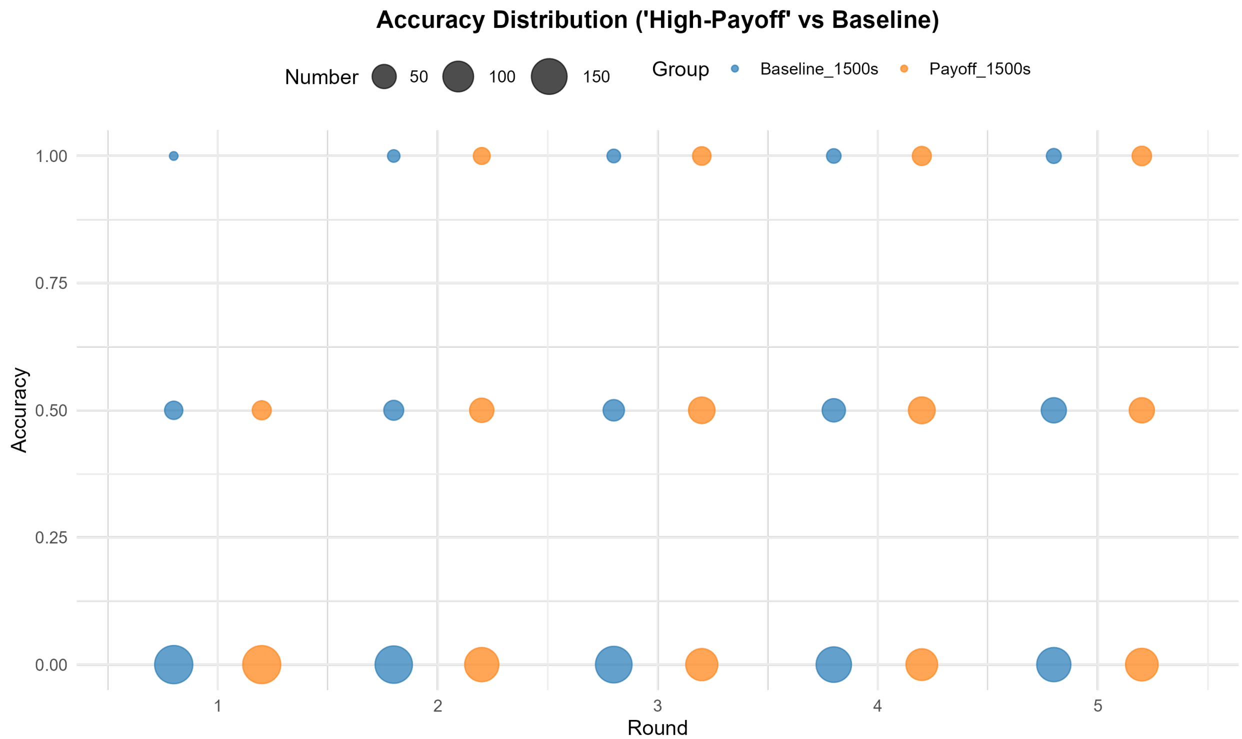
This mini-project provided valuable hands-on experience in modifying and interpreting a complex cognitive model. Implementing the WSLS, emotion feedback, and high-payoff strategy mechanisms required careful consideration of where and how to integrate these concepts within the existing evidence accumulation framework, reinforcing understanding of the model's dynamics. Analyzing the simulation outputs highlighted the non-trivial interactions between learning, decision heuristics, feedback effects, and strategic adaptation. Particularly insightful was the counter-intuitive finding for the 'Depressed' state and the clear dominance of the throughput-focused 'Direct High-Payoff' strategy, prompting reflection on task constraints and reward structures. Applying course content, particularly lectures on reinforcement learning (WSLS relevance), drift-diffusion models (core decision process), and cognitive control (strategy selection, feedback processing), was essential. The practical R programming and simulation analysis directly built upon workshop skills. For future projects, I could apply similar computational modelling techniques to investigate other cognitive biases or learning variations. The process demonstrated autonomy in translating conceptual ideas (like 'giving up') into specific model implementations and initiative in designing the comparative analysis of strategies. Potential action points include experimentally validating these strategic trade-offs in human participants and exploring more sophisticated meta-cognitive models for strategy selection.

Supplementary Figures



Supplementary Fig. 1: Behavioral metrics under the WSLS strategy.

Violin plots displaying the distribution of **a**, Schema Reaction Time (RT), **b**, Schema Observation Counts (OB), and **c**, Schema Attention Shifts (AS) during the schema identification phase. Comparisons are shown between the Baseline model and the Win-Stay, Lose-Shift (WSLS) model, with WSLS data further conditioned on the outcome of the preceding round: whether the same schema was selected as in the previous round and whether previous match was accurate (Same, AC=1), inaccurate (Same, AC<1), or if a different schema was selected (Diff, AC=1; Diff, AC<1). Horizontal lines within violins denote medians. Indicated pairwise comparisons in RT, OB, and AS show statistical significance between 'Same, AC=1' and baseline (P<0.001, determined by Wilcoxon rank-sum test following overall Kruskal-Wallis test).



Supplementary Fig. 2: Schema matching accuracy for high-reward strategies.

Bubble plot illustrating the frequency distribution of schema-matching accuracy across simulation rounds for agents employing the Baseline strategy (blue circles) versus the 'High-Payoff' strategy (orange circles). A two-sample Kolmogorov-Smirnov test indicated no significant difference between the overall accuracy (AC) distributions of the 'Direct High-Payoff' and Baseline groups (P>0.05).