Revisiting AI Rogers' Paradox: The Trade-Off Between Mastery Speed and Collective Innovation

Ghina Al Shdaifat (gha2009), Maitri Patel (mp7075), Jeevesh Attri (ja5593)

1 Motivation and Summary

Alan Rogers showed that cheap social learning offers no lasting fitness advantage over costly individual learning, a result known as Rogers' Paradox. More recently, *Revisiting Rogers' Paradox in the Context of Human-AI Interaction* introduced an AI agent that learns socially from humans, confirming the **AI Rogers' Paradox**: *the widespread availability of cheap AI social learning, even when trained on human data, may not improve society's collective world model in the long run.*

The original model treated knowledge as a binary state (adapted or not), which is useful for measuring accuracy but inadequate for capturing cumulative innovation or discovery of new knowledge.

We propose a multi-state **"Knowledge Map" or "Tech Tree"** to replace this binary setup. It contains two pathways: a **Low-Risk Branch** yielding quick gains but plateauing at a local minimum, and a **High-Risk Branch** with slower initial returns but leading to a global maximum.

Our **hypothesis** is that AI, by efficiently guiding users toward safe, low-risk knowledge (Mastery), may reduce incentives for the costly exploration required for true innovation (Discovery). We will test this trade-off by measuring both adaptation speed and the collective rate of new knowledge discovery.

2 Fit into Related Work

This project addresses several open questions raised by Collins et al.:

(a) **Enriched Realism:** Collins et al. suggested future work include "richer representations of our uncertain, dynamic world." Our Knowledge Map, where each agent's state is an n-value vector of discovered nodes, provides such realism. Nodes correspond to different survival probabilities, moving beyond the binary fitness metric.

(b) **Deskilling and Negative Feedback:** The original work showed that learning from AI reduces individual learning success (κ_j). We extend this with an incentive-based view: if the AI learns from the population and snaps to the **modal** or **average** state, it channels the population toward the Low-Risk branch. This easy optimization reduces incentives for the costly exploration needed to reach the global optimum.

(c) **Strategy Evaluation:** The source tested strategies such as Critical Social Learning (overriding AI output with individual learning) and Model-Centric AI updates. We will re-evaluate these within the Knowledge Map to see if they can restore balance between fast Low-Risk mastery and long-run High-Risk discovery.

3 Key Contributions and Hypothesis

Our primary contribution is the development and analysis of the Knowledge Map environment to measure the fundamental tension between exploitation and exploration in human-AI collective learning. We combine the metrics required by the original authors to test our hypothesis.

3.1 Architectural Contribution: The Knowledge Map

The environment is modeled as a multi-node, branching structure where agents progress incrementally through Individual Learning (IL) on the **same branch** or via Social/AI Learning (SL/AI) which allows movement across **different branches**.

Refinement: To address how the High-Risk branch is first discovered, we will add an **Exploratory IL action**. With a small probability, an agent performing IL may discover the initial node of the High-Risk branch instead of only advancing incrementally along the Low-Risk branch. In line with real-world uncertainty, agents will not know they are on a Low- vs. High-Risk path; they only observe immediate rewards and lack information about long-term potential. This framing makes the social dilemma more realistic.

3.2 Metrics and Primary Hypothesis

We track three interdependent metrics:

- **Mastery Score (Q_M):** The collective world understanding (average fitness value $V(K_j)$ of all agents' current nodes).
- **Discovery Rate (R_D):** The frequency with which agents reach a previously undiscovered knowledge node, specifically focusing on success along the High-Risk branch (Innovation).
- **Convergence Rate (T_c):** The number of iterations required for the Mastery Score (Q_M) to stabilize or reach a critical competence threshold (θ) (speed). Importantly, while slow convergence can preserve exploration, in volatile environments (high u) a **fast T_c may be adaptive**, allowing populations to lock onto a "good enough" solution before conditions shift. Thus T_c must be interpreted in conjunction with Q_M and u.

Primary Hypothesis: The availability of cheap AI social learning, which directs agents toward the modal/average state of the population, will significantly increase the Mastery Score and hasten Convergence (T_c) to the local optima (Low-Risk branch), but this gain in speed and short-term fitness will result in a decreased long-run Discovery Rate (R_D) of the High-Risk branch knowledge.

4 Plan for Implementation and Methodology

We will build on the open-source simulation code provided by Collins et al. Our plan is divided into three phases over 10 weeks.

4.1 Phase 1: Environment Implementation (Weeks 1-3)

Replicate the baseline model and implement the Knowledge Map. The aim here is to replace binary states with an n-value state vector, define the fitness landscape $V(K_j)$ for low- vs. high-risk branches, and update the learning mechanisms (IL, SL, AI).

1. **State Representation:** Implement the state vector $K_j(t)$ which replaces the binary adapted/not adapted status with an n-value corresponding to the agent's location on the Knowledge Map.
2. **Fitness Landscape:** Define the fitness function $V(K_j)$ such that nodes in the Low-Risk branch quickly increase survival probability (s_{OK}) but nodes in the High-Risk branch eventually lead to a higher maximum $V(K)$.
3. **Learning Rule Modification:**
   - **Individual Learning (IL):** By default, IL moves the agent incrementally to the next adjacent state on the same branch (exploitation). With a small probability, IL may instead trigger an **exploratory action** that discovers the initial node of the High-Risk branch.
   - **Social Learning (SL):** Instead of copying randomly from the entire population, SL will follow two refinements: (a) **Local SL**: agents may only copy peers "nearby" on the Knowledge Map, slowing down cross-branch diffusion. (b) **Status-Biased SL**: when sampling peers, agents give higher probability to those further along a branch (prestige bias) or with longer survival above the critical threshold (success bias). A tunable parameter α controls how strongly agents weight these high-status exemplars.
   - **AI Learning (AI):** The AI system snaps to the modal or average state of the population, thereby directing the agent to the most common (often Low-Risk) discovered knowledge.

4.2 Phase 2: Experimentation and Strategy Testing (Weeks 4-7)

Run large-scale simulations (T = 200,000 steps) under varying environmental volatility (u: probability of world change) and AI update schedules. We will test and compare four strategy conditions: (a) Individual Only baseline, (b) AI Social Only, (c) Critical Social Learning, and (d) Model-Centric update strategies.

- **Individual Only (Baseline):** Measures fundamental Discovery Rate equilibrium.
- **AI Social Only (Rogers' AI):** Agents learn only individually or socially from the AI (AI defaults to the modal state).
- **Critical Social Learning:** Agents engage in critical appraisal, switching to individual learning if the AI's output is deemed unsuccessful, testing if critical learning restores innovation.
- **Model-Centric:** Test the impact of varying the AI's update cadence (how frequently the AI learns from the population mean).

4.3 Phase 3: Analysis and Write-up (Weeks 9-10)

We will calculate and evaluate Mastery Score (Q_M), Discovery Rate (R_D), and Convergence Speed (T_c) for all simulation runs. The analysis will focus on the relationship between rapid convergence (T_c) to a local optimum (Mastery) and the corresponding suppression of the long-run Discovery Rate toward the global optimum. Results will explicitly interpret T_c in the context of environmental volatility (u) and path dependence driven by status-biased social learning.

References

[1] K. M. Collins, U. Bhatt, and I. Sucholutsky. Revisiting Rogers' Paradox in the Context of Human-AI Interaction. *arXiv preprint arXiv:2501.10476*, 2025. `https://doi.org/10.48550/arXiv.2501.10476`.

[2] K. M. Collins, I. Sucholutsky, U. Bhatt, K. Chandra, L. Wong, M. Lee, C. E. Zhang, T. Zhi-Xuan, M. Ho, V. Mansinghka, *et al.* Building machines that learn and think with people. *Nature Human Behaviour*, 8(10):1851–1863, 2024.

[3] M. Enquist, K. Eriksson, and S. Ghirlanda. Critical social learning: a solution to Rogers's paradox of nonadaptive culture. *American Anthropologist*, 109(4):727–734, 2007.

[4] A. R. Rogers. Does biology constrain culture? *American Anthropologist*, 90(4):819–831, 1988.