UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Superstitious Ordering Despite Discouragement: The Role of Visual Features

Permalink

https://escholarship.org/uc/item/1pb3175x

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Cai, Linzhuohao
Gong, Tianwei
Lagnado, David

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Superstitious Ordering Despite Discouragement: The Role of Visual Features

Linzhuohao Cai (linzhuohao.cai.24@alumni.ucl.ac.uk)

Tianwei Gong (t-gong@ucl.ac.uk)

David Lagnado (d.lagnado@ucl.ac.uk)
Department of Experimental Psychology, University College London, 26 Bedford Way, London, WC1H 0AP, UK

Abstract

Humans readily perceive patterns even in situations with only random noise, but how do they respond when the genuine rule deliberately discourages this tendency? We investigated superstitious ordering under a consistent preference-discouraging (PD) schedule, in which reward probability increases for previously avoided options and decreases for previously chosen options. This contradicts typical reinforcement principles, in which some options are always preferred over others. We also examined the role of visual features, with one group using geometric shape combinations as stimuli and another group using natural images. We found that participants in both groups developed significant subjective linear orderings of stimuli, with a marginally stronger effect in the Geometrical Image condition. Comparisons with the benchmark agent, optimized Q-learning agent, and random agent showed that none of the models could explain participants' response patterns, indicating a unique and stable superstitious ordering effect in human learning.

Keywords: superstitious learning; randomness; visual feature; transitive inference.

Introduction

Humans possess a fundamental drive to detect patterns, a mechanism that facilitates learning but can also lead to "superstitious behavior", which refers to the formation of spurious beliefs about causal relationships where none exist (Skinner, 1948; Matute et al., 2015). While superstition is typically studied under random reinforcement (Skinner, 1948), a stricter test arises in Preference-Discouraging (PD) schedules. We adopt the transitive inference paradigm and will focus on it in this paper. In a transitive inference paradigm, each trial asks participants to choose between two stimuli from a set. The standard procedure ranks all stimuli and aims to examine whether people could learn the full orderings even if they don't see the feedback of all possible pairs (Jensen et al., 2021). The random procedure gives participants random feedback, while a PD procedure prescribes that reward probability increases for previously avoided options and decreases for previously preferred options (Jin et al., 2022).

Although the PD environment imposes a simple rule – to alternate between options as much as possible instead of consistently preferring some options over others – it requires a counterintuitive "win-shift" mindset which contradicts the common "win-stay" heuristic (Bonawitz et al., 2014; Nowak & Sigmund, 1993). Consequently, learners failing to grasp the dynamic rule may default to a simpler, albeit incorrect, strategy: assuming a static hierarchy among stimuli (e.g., A > B > C).

Recent work suggests that monkeys spontaneously develop such subjective orderings in random stimuli as well as in PD stimuli (Jin et al., 2022). Repeating the same study with humans revealed greater individual differences rather than a coherent pattern (Jin et al., 2024): some people formed subjective orderings in both a random set and a standardly ranked set, whereas others did not form subjective orderings in either. With regard to the PD procedure, some participants formed subjective orderings in both a PD set and a standardly ranked set, some formed them only in the standardly ranked set, and some formed none in either set. One major feature of these studies was that learners were required to learn two sets simultaneously, one ranked and one unranked (i.e. rewarded randomly or under the PD rule). In this context, the presence of a ranked set may encourage participants to assume that all stimuli are ranked. It remains unclear what happens in a learning environment in which none of the stimuli have inherent ranks. Therefore, here we test whether people will form subjective, superstitious orderings in a pure PD task that excludes any explicitly learnable rankings.

We also examine the role of visual features in this study. We hypothesize that more salient and easily comparable visual features facilitate the formation of superstitious rankings. Given the true PD rule is counterintuitive, learners may mistakenly attribute rewards to rules related to visual characteristics rather than to choice history. If the visual characteristics of stimuli are easier to extract and compare to each other, this may encourage people to remain in a rank-learning mindset. Conversely, in the absence of easily comparable visual characteristics, learners may abandon this mindset, instead perceiving the environment as random or, potentially, becoming more likely to uncover the true PD rule. Therefore, we manipulated the stimulus format such that half of the participants engaged with geometric shapes, while the other half viewed natural images (Figure 1).

Crucially, we compared human performance against three computational models: (1) a Random Response Agent (baseline), (2) a Benchmark Rational Agent (who knows the

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society* ©2026 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

underlying PD rule and acts to gain the highest reward), and (3) Optimized Q-learning Agents (standard model-free reinforcement learning). These comparisons allow us to explore (1) whether people form subjective orderings (compared with a random agent); (2) whether people in certain conditions can discover the true rule (compared with a benchmark agent); and whether any human superstitious ordering is driven by associative learning (mimicking the Q-learner).

Subjective Ordering Analysis

Standard metrics of learning, such as total reward or simple choice proportions, are insufficient for detecting superstitious behavior in PD tasks where no ground-truth ordering exists. A participant might choose consistently (high bias) or randomly (low bias), but these aggregate statistics fail to capture the structural consistency of their preferences. To address this, we employed Subjective Ordering Analysis (SOA), a Bayesian framework that infers latent preference hierarchies from trial-by-trial choices (Jensen et al., 2019; Jin et al., 2022).

The core intuition of SOA is that agents represent the value of each stimulus not as a fixed number, but as a probability distribution on a latent "mental number line" defined by a mean position μ and uncertainty σ (see Methods). In this framework, "superstition" is strictly defined as the emergence of a stable, transitive hierarchy among stimuli whose visual identities are orthogonal to the reward structure. By fitting this model, we can distinguish the *strength* of their subjective preferences from simple random guessing (see Results). This allows us to quantify the degree to which an agent imposes a structured reality onto a stochastic environment.

Method

Participants

Ninety participants were recruited via Prolific. Three were excluded due to data transmission errors, resulting in a final sample of 87 participants. Participants were randomly assigned to one of two conditions. The Geometrical Image condition included 44 participants (26 male, 17 female, 1 non-binary; M_{age} = 41, range: 23 - 67), while the Natural Image condition included 43 participants (23 male, 20 female; M_{age} = 44, range: 19 - 70). All participants provided informed consent and received a base payment of 1.5 pounds plus a performance-based bonus (see Task Design). An experiment demo is available online (https://github.com/Thredson/exp2025)

Stimuli

Stimuli consisted of six distinct images per condition (Figure 1a). In the Geometrical Image condition, the images were geometric shapes that varied systematically along clearly definable dimensions (e.g., number, color, size, position), providing an easy-to-access feature base for comparisons between stimuli. These stimuli were designed based on the children's game *Zendo*. Zendo has been used to study complex concept learning because it affords an infinite number of possible concepts within the hypothesis space generated by combinations of dimensions (e.g., Bramley et al., 2018; Fränken et al., 2022). This property is important for the present study because it provides participants with countless ways to hypothesize the underlying basis for how the stimuli are ranked.

In the Natural Image condition, stimuli were natural images (e.g., heterogeneous scenes) that were visually discriminable but lacked comparable systematic geometric regularities (see examples in Figure 1 below). The images were adopted directly from Jin et al. (2022) and have been shown to trigger subjective orderings.

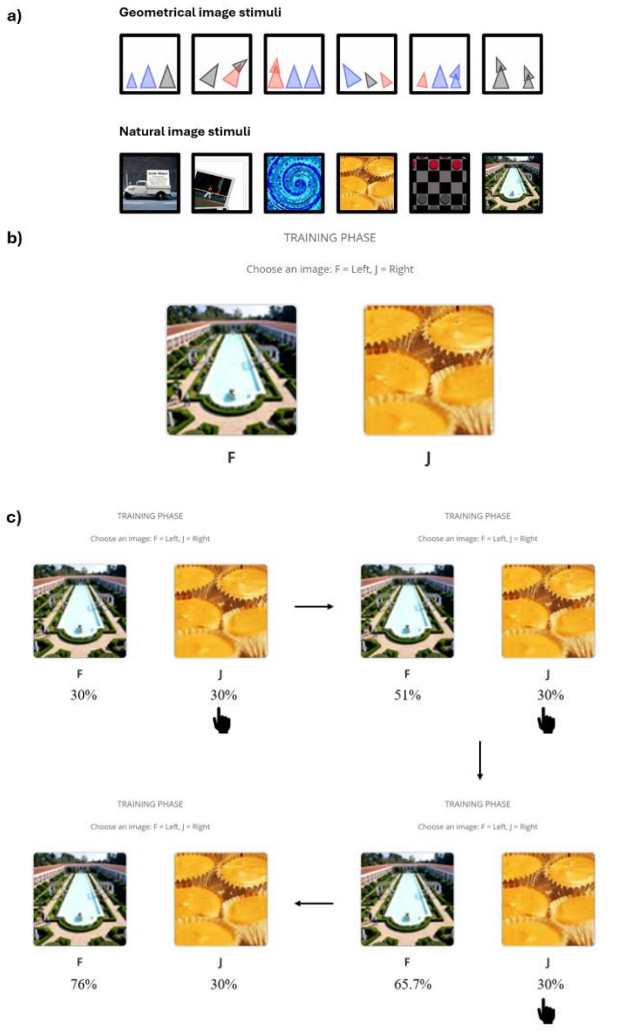

Figure 1. a) Images used in both conditions. b) The response interface where a pair of stimuli was displayed for the participant to select from during the experiment. c) An example of the change of reward probability of an option that is being consistently ignored.

Task Design

We employed a pure Preference-Discouraging (PD) procedure adapted from Jin et al. (2022). In a sequence of 140 trials, participants chose between two images from the stimulus set (Figure 1b). Crucially, the reward probability depended entirely on choice history, following a principle that penalizes consistent preferences. Specifically, the reward probability for a stimulus was calculated as:

$$P(\text{reward}) = 1 - (1 - 0.3)^{(n+1)}$$

where *n* represents the number of consecutive trials the option was available but not chosen. *P(reward)* increases as *n* increases. This schedule punishes "win-stay" behavior and encourages an alternation between choices. Every time people got rewards, they got a 1 penny bonus.

Procedure

The experiment was programmed using jsPsych. Participants were instructed that their goal was to discover how to maximize their total bonus earnings given a choice making task. They were explicitly told about receiving either the exact reward of 1 penny or no reward after each trial, but they were not informed of the specific history-dependent nature of the task.

The task consisted of 140 trials divided into two phases: a Training Phase (Trials 1–80) and a Testing Phase (Trials 81–140). Images were labeled A–F in the programming code, but these labels were used only for later analysis and were not visible to participants. In the Training Phase, participants chose between pairs of stimuli drawn from pseudo-transitive pairs (AB, BC, CD, DE, EF) along with their reversed versions (BA, CB, DC, ED, FE). This led to 10 unique trials which was counted as a block. Participants went through 8 training blocks (80 trials). Upon making a choice, they received immediate feedback indicating whether they won a reward or not, and their accumulated bonus was updated. In the Testing Phase, stimuli consisted of all possible 15 pairs (AB, AC, AD …) along with their reversed versions which consists of 30 pairs. Participants went through two testing blocks (60 trials). The feedback was withheld in the testing phase to assess the stability of learned preferences without further reinforcement signals. However, participants were informed in the instructions and knew that rewards continued to accumulate in the back.

In each trial, two images were presented on the left and right sides of the screen. Participants selected the left image by pressing the "F" key and the right image by pressing the "J" key. There was no time limit for the response. The order of trials was randomized within each, and labels of each image were randomized between participants.

After completing the task, participants completed a debriefing survey assessing their explicit knowledge of the reward structure. They were asked, "Which of the following best describes the rule for getting rewards?" and selected from four options: (1) "Pictures with some specific features should be preferred" (Feature-dependent); (2) "The picture you picked previously should not be picked on the next occasion, which means the picture not picked previously should now be picked" (PD rule, which was the ground truth); (3) "There is a ranking between six pictures. For example, Picture 1 > Picture 2 > Picture 3... A picture with a relative higher rank in the present pair should be preferred." (Fixed Ranks); (4) "Completely random" (Random). Finally, demographic information was collected before participants received their payment code.

Computational Benchmarks

To evaluate human performance, we simulated three different agent types on the exact same task sequences. For each model, we simulated 200 agents to compute the average reward, providing a robust estimate of performance levels. We subsampled 44 agents (matching the human sample size per condition) to conduct the computationally intensive Subjective Ordering Analysis. The agents were defined as follows:

Random Agent Selected two presented options with equal probability ($P = 0.5$).

Benchmark Rational Greedy Agent A greedy algorithm that knew the ground truth rule (see Task Design) and deterministically selected the option with the higher current *P*(reward).

Optimized Q-learning Agent A standard model-free reinforcement learning agent (Sutton & Barto, 2018). It updated action values via $Q(s) \leftarrow Q(s) + \alpha[r - Q(s)]$ and selected choices using a Softmax policy:

$$P(\text{choosing option i}) = \frac{e^{\left(\frac{Q_i}{\tau}\right)}}{\sum e^{\left(\frac{Q_j}{\tau}\right)}}$$

We performed a grid search over α in [0.1, 0.9] and τ in [0.5, 3.0]. The best-performing parameters were α = 0.2 and τ = 3.0.

Subjective Ordering Analysis (SOA)

To quantify superstitious preferences, we employed Subjective Ordering Analysis (Jensen et al., 2019). We fit a hierarchical Bayesian model assuming each stimulus *i* occupies a latent position z_i with uncertainty σ_i. The probability of choosing Right over Left is modeled as:

$$\text{P(choose right)} = \frac{\theta}{2} + (1 - \theta) \times \phi\left(\frac{z_{right} - z_{left}}{\sqrt{\sigma_{left}^2 + \sigma_{right}^2}}\right)$$

where θ is a lapse parameter and φ is the standard normal cumulative distribution function. We extracted posterior

mean positions to calculate the Subjective Ordering Slope, where a more negative slope indicates a stronger hierarchy.

Results

Our analysis proceeded in four stages: (1) examining explicit conscious beliefs about the task rule; (2) quantifying implicit superstitious behavior using the Subjective Ordering Analysis (SOA); (3) comparing human performance with computational agents; and (4) analyzing reward acquisition of each group. A repository of analysis script is available online (https://github.com/Thredson/exp2025).

Conscious Understanding of Task Structure

Post-experiment surveys asked participants to identify the rule governing rewards. As shown in Figure 2, the answer distribution of two conditions differed. A chi-square test of independence confirmed that the distribution of perceived patterns differed marginally between conditions ($\chi^2(3) = 7.58$, $p = 0.056$). In the Geometrical Image condition, participants were numerically more likely to attribute rewards to intrinsic visual properties rather than choice history. Specifically, 17 participants (38.6%) in the Geometrical Image condition believed the rule was feature-dependent, compared to only 9 participants (20.9%) in the Natural Image condition. Crucially, the presence of geometrical features appeared to mask the true rule. Only 3 participants (6.8%) in the Geometrical Image condition correctly identified the PD rule. In contrast, more participants (11, 25.6%) in the Natural Image condition were able to choose the ground truth.

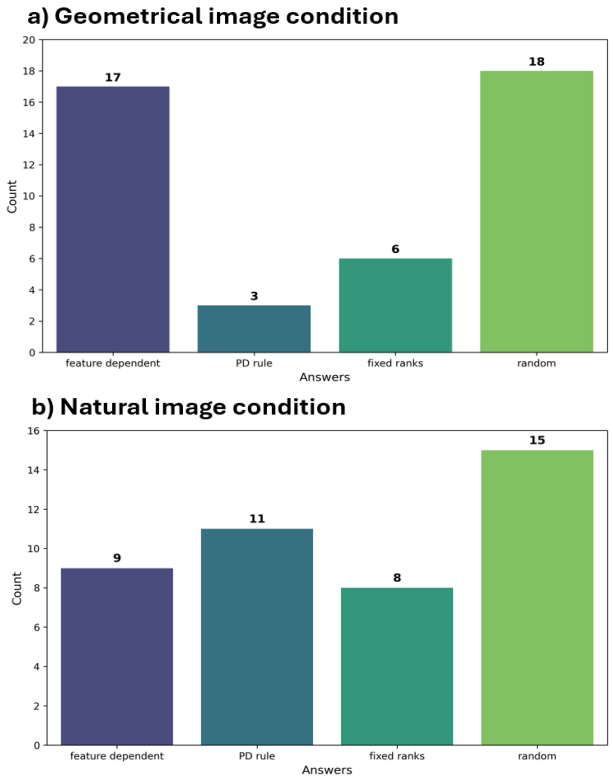

Figure 2. Distribution of self-reported task understanding. a) Participants in the Geometrical Image condition frequently defaulted to feature-based rules (dark blue bar). b) Participants in the Natural Image condition were less likely to focus on features and more likely to identify the true "Reverse" pattern or perceive the task as random.

The Emergence of Implicit Superstition

Moving beyond explicit reports, we used Subjective Ordering Analysis (SOA) to quantify implicit behavioral biases. Figure 3 illustrates the posterior distributions of latent stimulus positions for a representative participant. For a perfectly rational or random agent, these distributions should overlap heavily, reflecting the functional equivalence of the stimuli. However, the data from this participant reveal clear, non-overlapping distributions. For instance, Stimulus D is consistently separated from Stimulus C in the posterior space. This separation indicates a high-certainty belief in a false hierarchy, despite the environment offering no statistical support for such distinctions.

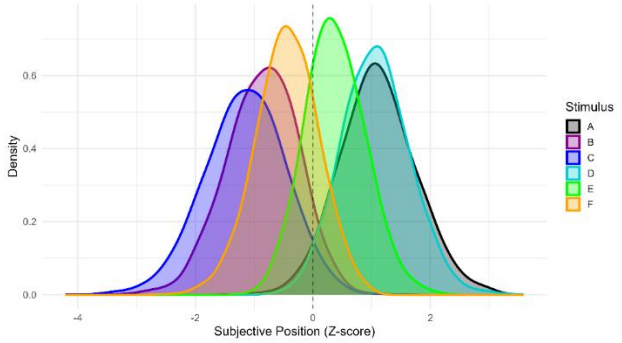

Figure 3. Representative SOA density plot for a single participant. The distinct, non-overlapping posterior distributions demonstrate that the participant learned a stable, transitive hierarchy among stimuli, despite all stimuli having identical reward functions.

To quantify this effect at the group level, we calculated a Subjective Ordering Slope for each participant by fitting a linear regression that predicts the exact subjective positions (z-scores in Figure 3) from subjective ranks (1-6). If six z-scores are close to each other (the learner treats six stimuli as equally), the slope is near zero; if they are well separated, the slope is significantly below 0, indicating superstitious ordering.

One-sample t-tests against zero confirmed that superstition was robust and pervasive. Participants in the Geometrical Image condition exhibited strong negative slopes ($M = -0.35$, $SD = 0.13$), significantly different from zero ($t(43) = -15.73$, $p < 0.001$, $BF_{10} = 8.05 \times 10^{17}$). Similarly, participants in the Natural Image condition also developed significant orderings ($M = -0.29$, $SD = 0.15$, $t(42) = -11.10$, $p < 0.001$, $BF_{10} = 4.57 \times 10^{12}$).

Critically, we compared the strength of this bias between groups. An independent samples t-test revealed a marginally significant difference ($t(85) = -1.89$, $p = 0.062$, $d = 0.43$, $BF_{10} = 1.07$). This provides a weak indication that salient comparable features may further facilitate the formation of subjective orderings.

Humans and Computational Agents

We compared human SOA slopes against our three computational benchmarks: the Random Agent, the Benchmark Rational Agent, and the Optimized Q-learning Agent (Figure 4). This comparison can further test whether human behavior could be explained by simple associative learning (Q-learning) or random exploration.

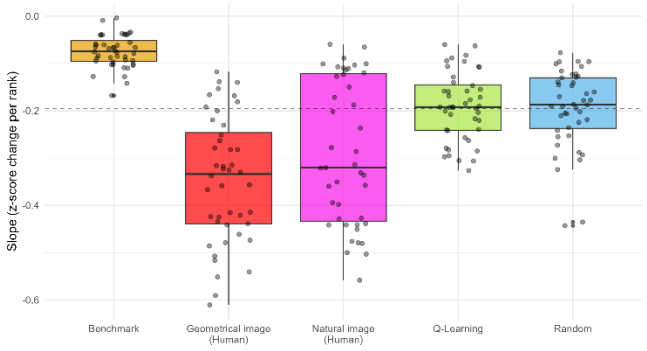

Figure 4. Comparison of Ordering Slopes across Humans and Agents. The Benchmark, Q-learning, and Random agents cluster near zero (indicating no stable hierarchy). In contrast, Humans in both conditions show strong negative slopes, indicating the imposition of superstitious structure. The dissociation highlights that human superstition is not explained by standard RL mechanisms. The horizontal dashed line represents the mean slope of Random agents.

We first demonstrated the results of sampled agents. As expected, the Benchmark Rational Agent showed no ordering bias (M = -0.1 from sampled agents) since it optimally tracked the dynamic reward probabilities. Similarly, the Random Agent produced a small slope of M = -0.17, representing the noise floor of the regression analysis.

Crucially, the Optimized Q-learning Agent also exhibited slopes near the noise floor (M = -0.19). The high optimal temperature ($\tau = 3.0$) indicates that for a standard RL agent, the best strategy in this PD task is to behave almost randomly to avoid the "win-stay" penalty. Consequently, the Q-learner "learned to be random" and did not form stable hierarchies.

Humans, however, stood apart from all these models. Figure 4 reveals a qualitative dissociation: while all agents cluster near the zero/noise baseline, human distributions appear to be negative. We performed contrasts to verify this dissociation:

Humans vs. Random To confirm that human ordering was not merely statistical noise, we compared each human group against the Random Agent baseline (a slope of -0.19). Participants in the Geometrical Image condition exhibited significantly more negative slopes than the Random Agent ($t(43) = -6.88$, $p < 0.001$). Similarly, the Natural Image condition also showed significantly more negative slopes than the Random baseline ($t(42) = -3.97$, $p < 0.001$). This confirms that behavior in both conditions was systematically structured, distinct from stochastic responding.

Humans vs. Q-learning We tested whether human superstition could be explained by model-free reinforcement learning. The Geometrical Image condition slopes were significantly more negative than the Optimized Q-learning Agent ($t(43) = -7.08$, $p < 0.001$). Likewise, the Natural Image condition was significantly more negative than the Q-learning Agent ($t(42) = -4.14$, $p < 0.001$).

This result challenges the assumption that superstition in this context is a byproduct of model-free reinforcement learning errors. While the Q-learning agent learned to be random to maximize rewards, humans learned a complex, structured, but objectively false hierarchy. The dissociation suggests that human superstition involves a distinct cognitive drive to impose structure that overrides both rationality and simple reward-based reinforcement.

Reward Efficiency and the Cost of Superstition

Finally, we asked whether this imposition of structure conferred any advantage or cost. We compared the average reward rates earned by humans against the computational agents. The Optimized Q-learning agent achieved an average reward rate of 45.4% (derived from its random-like strategy). Moreover, the Benchmark Agent gained 49.15% rewards (the benchmark agent here is based on a greedy policy, which may limit the upper bound of the achievable reward, for which we will further discuss in the Discussion). By comparison, the Random Agent baseline was 45.7%. Human participants, despite their elaborate ordering strategies, achieved an average reward rate of 43.8% across conditions.

A one-sample t-test comparing human performance against the Optimized Q-learning benchmark revealed that human performance was similar to Q-learning agents ($t(43) = -0.89$, $p > 0.05$). In fact, human performance was statistically indistinguishable from the Random Agent ($t(43) = -1.45$, $p > 0.05$). However, human performance was significantly different compared to Benchmark Agent ($t(43) = -7.71$, $p < 0.001$). This result highlights the maladaptive nature of the observed superstition. By adhering to a rigid hierarchy, participants systematically over-sampled their subjectively preferred options.

General Discussion

The present study investigated whether humans spontaneously form superstitious preference hierarchies in a pure

Preference-Discouraging (PD) environment and whether salient visual features exacerbate this tendency. Our results provide converging evidence that humans impose structured orderings onto stimuli when the optimal strategy is counterintuitive. Specifically, we found that: (1) participants in both Geometrical and Natural Image conditions developed significant subjective linear orderings (negative slopes) that were significantly stronger than chance; (2) results indicate that how visual features are presented to participants (easily comparable geometric images vs. complex natural images) may influence the beliefs they form and their behaviors, although whether this effect is robust requires further investigation; (3) crucially, human behavior qualitatively dissociated from Optimized Q-learning agents, which learned to be random in order to mitigate the PD penalties.

Implications

The dissociation between human and agent performance challenges the view of superstition as a mere associative error. While Q-learning agents optimized by learning to be random, humans appeared to operate under a strong structure-learning prior, imposing complex, transitive hierarchies where none existed. This cognitive drive proved maladaptive: by adhering to rigid preferences, participants systematically oversampled their subjectively preferred options, thereby degrading their reward probabilities under this environment. Consequently, our results suggest that the human cognitive system exhibits a strong prior for pattern seeking, often prioritizing structural coherence over reward maximization even in environments that specifically punish such behavior.

Comparison with Prior Work

Our findings extend previous research on superstitious ordering in key ways. Jin et al. (2022) demonstrated that monkeys form subjective orderings in PD tasks, and later work in humans (Jin et al., 2024) found similar effects. However, these prior studies often employed mixed blocks containing both PD (unranked) and learnable (ranked) pairs. In such contexts, the spillover of a ranking mindset from learnable trials could explain the superstition. The present study employed a pure PD task with no ground-truth rankings. The persistence of strong ordering effects here suggests that the drive to hierarchically organize stimuli is intrinsic to the learner, not merely an artifact of task context. Furthermore, our manipulation of visual features adds a new dimension. Humans have been shown to have strong concept and category learning abilities (Carey, 2009), and prior research has demonstrated that people can easily learn the order between stimuli when these orders are based on underlying categories (Jensen et al., 2021). Here, we consider whether different stimuli can be readily decomposed into comparable dimensions and how this may affect the results. Future research could further explore this topic to investigate the relationship between basic concept and category learning, the ranking mindset, and other superstitious beliefs. Third, we added questionnaires to explicitly assess participants' beliefs about the ground truth. This approach helps identify potential differences between conditions and can also serve as an additional measure of participants' beliefs alongside behavioral measures.

Limitations and Future Directions

Several methodological limitations warrant discussion and guide future research. First, we detected a marginal difference between conditions with different visual features in both explicit belief reports and behavioral measures. However, the robustness of this effect may require a larger sample size to confirm.

Second, the difference between the Benchmark Rational Agent and the Random Agent is relatively small. This may limit participants' motivation to learn the true rule. Future research should adjust the reward structure to ensure a larger gap between benchmark and random performance. It may also be useful to increase the rewards associated with the Random Agent to examine whether participants could perform even worse than random due to superstitious ordering (Jin et al., 2022).

Third, the current benchmark is based on a greedy strategy. Future research could implement agents with different policies (e.g., model-based planners looking k steps ahead or Recurrent Neural Network (RNN) that rely on internal memory) to see if a higher theoretical upper bound of performance exists and to further contextualize the cost of human superstition.

Conclusion

This study demonstrates that superstitious behavior in preference-discouraging environments is not merely a failure of learning, but a consequence of an active cognitive drive to find order. Humans spontaneously construct hierarchies among equal options, a tendency that is distinct from simple associative learning and potentially exacerbated by visual saliency. This suggests that the cognitive drive for order operates as a rigid constraint: while advantageous for reducing dimensionality, this mechanism limits behavioral flexibility, leading to the maintenance of suboptimal policies in preference-discouraging contexts.

Acknowledgment

Linzhuohao Cai was supported by the dissertation research funding in UCL Department of Psychology. Tianwei Gong was supported by the British Academy Postdoctoral Fellowship Scheme (PFSS24\24009). We thank Victor Btesh for his very helpful discussions in the initial stage of this project.

References

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-Stay, Lose-Sample: A simple sequential algorithm for approximating Bayesian inference. *Cognitive Psychology*, 74, 35-65.

Bramley, N. R., Rothe, A., Tenenbaum, J. B., Xu, F., & Gureckis, T. M. (2018). Grounding compositional hypothesis generation in specific instances. *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, 1390–1395.

Carey, S. (2009). The origin of concepts. Oxford University Press.

Fränken, J. P., Theodoropoulos, N. C., & Bramley, N. R. (2022). Algorithms of adaptation in inductive inference. *Cognitive Psychology*, 137, 101506.

Jensen, G., Kao, T., Michaelcheck, C., Borge, S. S., Ferrera, V. P., & Terrace, H. S. (2021). Category learning in a transitive inference paradigm. *Memory & Cognition*, 49(5), 1020-1035.

Jensen, G., Muñoz, F., Alkan, Y., Ferrera, V. P., & Terrace, H. S. (2019). Reward associations do not explain transitive inference performance in monkeys. *Science Advances*, 5(7), eaaw2089.

Jin, Y., Jensen, G., Gottlieb, J., & Ferrera, V. (2022). Superstitious learning of abstract order from random reinforcement. *Proceedings of the National Academy of Sciences*, 119(35), e2202789119.

Jin, Y., Jensen, G., Gottlieb, J., & Ferrera, V. (2024). Learning of predictable rules mixed with random reinforcement in humans. *2024 Conference on Cognitive Computational Neuroscience*.https://2024.ccneuro.org/pdf/318_Paper_authored_CCN-abstract-YJ.pdf

Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A., & Barberia, I. (2015). Illusions of causality: How they bias our everyday thinking and how they could be reduced. *Frontiers in Psychology*, 6, 888.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 364(6432), 56-58.

Skinner, B. F. (1948). “Superstition” in the pigeon. *Journal of Experimental Psychology*, 38(2), 168-172.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.