Distinguishing Permanent from Transitory Changes: Experimental Evidence on Human Forecasting Ability*

Elise Payzan-LeNestour† Yunshen Yang‡ Qihe Tang§

February 4, 2025

Abstract

Decision-makers face a universal challenge, one particularly visible in financial investing: teasing apart whether changes are transitory disruptions or permanent structural shifts under deep uncertainty about shock nature. In a preregistered lab experiment, we examine how people navigate this cognitive challenge and explore methods to enhance their performance. Participants, incentivized with significant monetary rewards, had to forecast shifts in an asset value by first extracting an underlying pattern, a crucial prerequisite for successful forecasting by design. Participants outperformed expectations, demonstrating surprising pattern recognition abilities. A second group, merely informed of the pattern's existence without explanation, performed comparably to a fully informed control group, suggesting minimal guidance sufficed to enhance forecasting performance. The findings support the "active thinker" view of human decision-making and suggest potential for human-AI collaboration in finance, combining human cognition with AI's execution capabilities.

Keywords: Forecasting; Mental models; Information nudges; Bounded rationality; Behavioral economics; Experimental economics; Machine psychology

1 Introduction

Asset returns are known to intermingle fundamental changes in market dynamics ("regime shifts") with transitory, immediately-reversing fluctuations ("aberrations").[1] Real-time differentiation be-

*We thank Tony Berrada, Peter Bossaerts, Mark Dean, Etienne Koechlin, Mael Lebreton, John List, John Quiggin, Edward Shore, Susan Thorp, and seminar participants at the University of Chicago, Geneva University, Vienna University of Economics and Business, the University of New South Wales School of Economics, the University of New South Wales School of Banking and Finance, and the University of New South Wales School of Risk & Actuarial Studies, for very helpful comments. Funding from the University of New South Wales Business School is gratefully acknowledged.

†University of New South Wales Business School, elise@elisepayzan.com

‡University of New South Wales Business School, yunshen.yang@unsw.edu.au

§University of New South Wales Business School, qihe.tang@unsw.edu.au

[1]For example, Mandelbrot (1963), Embrechts et al. (2013), and d'Acremont and Bossaerts (2016).

tween these two types of shocks is crucial, as they demand opposing responses: regime shifts necessitate rapid strategic adaptation, while aberrations require restraint—maintaining the current strategy. Failure to distinguish between the two can lead to detrimental outcomes at both individual and systemic levels.[2] This challenge faced by investors exemplifies a ubiquitous cognitive task that transcends both scale and domain, from individual households to policymakers, from economics to ecology and geology: teasing apart temporary disruptions from permanent structural changes. The recent Covid-19 shock is a case in point, as all of us grappled with deep uncertainty about the shock's duration (Acharya et al., 2023). Importantly, such uncertainty may be mitigated when structural changes tend to be preceded by measurable patterns. For example, heightened microseismic activity and changes in gas emissions may be harbingers of major structural shifts such as volcanic eruptions, much as early warning economic signals may serve as reliable precursors to economic recessions. Thus, by uncovering these underlying pre-shift patterns, one could in principle distinguish structural changes from temporary disruptions.

Despite its importance, we know little about people's capacity to navigate this fundamental challenge in practice. Furthermore, the potential of behavioral interventions—such as "information nudges" (Sunstein, 2016) or "boosts" (Hertwig, 2017)—to enhance performance in this domain, also remains largely unexplored.

This paper addresses both questions through a controlled laboratory experiment. Participants engaged in a financial forecasting task with real and significant monetary incentives. Optimal performance required uncovering a visual pattern to forecast regime shifts in an asset value unfolding in real time. By design, asset value shifts were intermixed with aberrations (Fig. 1). Differentiating between the two types of shocks necessitated tracking the value of an indicator intrinsically linked to the asset value by a pattern. Critically, task participants were not informed about the existence of this pattern in the pre-task instructions. Consequently, they faced deep uncertainty—also known as structural or "Knightian" uncertainty (Knight, 1921), defined as "an uncertainty about the environment rather than an uncertainty in the environment" (Packard et al., 2021). That is, participants knew neither the key task parameters (aberration and shift probabilities) nor the Data Generating Process (DGP) itself—how events were generated and what defined the relevant contexts. We assessed forecasting accuracy by recording participants' trial-by-trial forecasts over 200 periods.

Success in the task requires participants to clear two high bars. First, they must extract the pattern linking indicator and asset values—a pattern that is "complex" in the precise sense defined by Kendall and Oprea (2024) (we elaborate below). Second, they must estimate switch probability using this pattern, itself a notoriously challenging task for human subjects (Benjamin,

[2]For instance, the 2010 "flash crash" alone erased $1 trillion in market value.

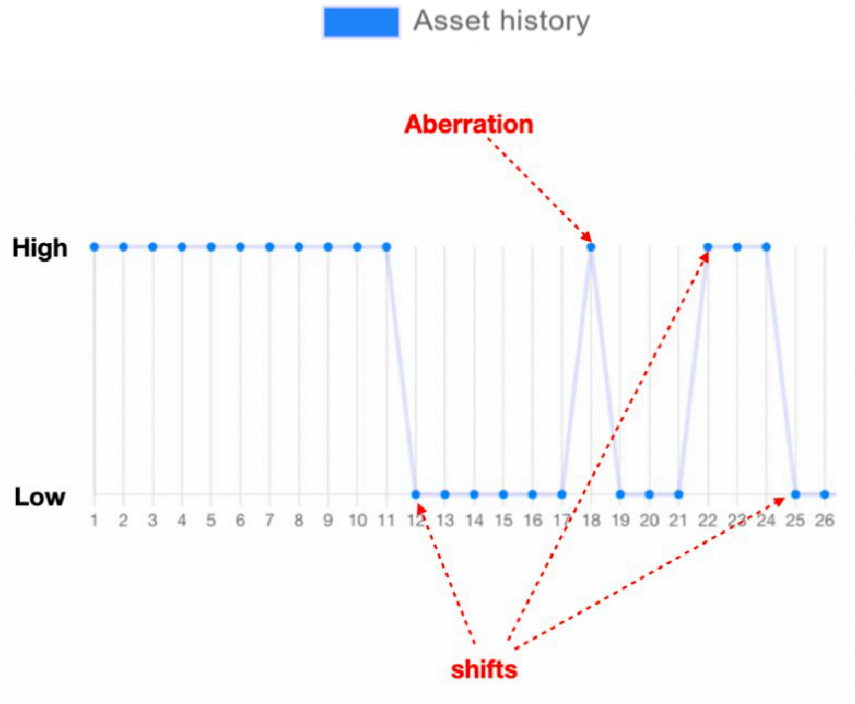


Figure 1. Simulating shifts and aberrations in a laboratory setting. The experimental task involves visualizing a financial asset's time series, constructed via a synthetic process. In this process, the asset value randomly switches between 'High' and 'Low' states. Some switches represent regime shifts (where the new asset value usually persists for several consecutive trials, e.g., shift from High to Low in trial 12), while others constitute purely transitory shocks ("aberrations," where the asset value reverts immediately, e.g., trial 18). The chart evolves in real-time—similar to those in computer trading platforms—plotting the newly realized asset value at each trial along with trailing values over a fixed window of recent past. Participants are tasked with forecasting the next asset value in each of 200 trials.

2019). Given this double hurdle, to give optimal behavior a chance of happening in our task, we provided substantial monetary incentives for optimal performance, following the established view in experimental economics that such incentives are necessary to measure cognitive capabilities in demanding tasks.[3] Without proper incentives, failure to observe optimal forecasting could reflect lack of motivation rather than cognitive limitations. Substantial incentives are thus crucial here for power—our ability to detect genuine effects. Beyond ensuring power, they also serve ecological relevance by mimicking real-world stakes as much as possible. These considerations, along with analytical tractability, were key criteria guiding our experimental design (detailed in Section 2.1).

Given the pattern's complexity, we hypothesized that participants would fail to extract it and consequently resort to "*model-free*" learning—updating values based on realized rather than expected outcomes, thus bypassing the need to understand the DGP. This characterization of participant behavior (detailed in Section 2.2) draws on the classic distinction between probabilistic

[3]For example, Wilcox (1993), Hertwig and Ortmann (2001), and Oprea (2024).

inference ("model-based") and adaptive learning ("model-free").[4]

We document a striking result: our hypothesis is falsified. The data reveal that a majority of participants successfully extracted the pattern and subsequently learned shift probability.

Turning to the question of whether forecasting performance can be boosted by an information nudge, we conducted a second experimental treatment that replicated the conditions of the original experiment, except for one key difference: the task instructions revealed the pattern's presence to the participants, requiring them to identify it rather than uncover it from scratch.

We find that participants in this second treatment demonstrated enhanced pattern recognition and performed comparably to a third (control) group that received full disclosure of the task structure in the task instructions, showing the effectiveness of our basic nudge. The high degree of control afforded by the laboratory experimental approach allowed us to isolate this effect of nudging.

Finally, we document an interesting asymmetry in participant behavior: while they successfully navigated regime shifts, they tended to mishandle aberrations, treating these transitory shocks as if they were shifts. Strikingly, this overreaction persisted even with the control group, which was explicitly informed about the purely transitory nature of aberrations. Taken together, these findings point to a fundamental duality in human cognition: individuals can handle structural changes in their environment, even under deep uncertainty (first treatment), yet systematically overreact to transitory shocks—even when explicitly informed of their purely transitory nature (control treatment).

Relation to literature

This study contributes to a rich forecasting literature, where earlier findings mirror the cognitive duality exhibited by our task participants—successful adaptation to regime shifts contrasting with overreaction to purely transitory shocks—and provide insights into its root causes. Afrouz et al. (2023) show how recency bias in AR(1) forecasting triggers overreaction when persistence is low. Closest to our study, Kremer et al. (2011) document intensified belief updating in response to large prediction errors, whether caused by transitory shocks or regime shifts under deep uncertainty about the shock nature. This same intensified response generates both overreaction in stable noisy environments (referred to by Kremer et al. (2011) as "*system neglect*") and suitable adaptation in changing environments, paralleling the duality observed in our task participants. Our experimental design marks a clear departure from this prior work by introducing novel features—notably an

[4]For example, Pouget (2007), Alós-Ferrer and Garagnani (2023), and Barberis and Jin (2023). Evidence suggests the human brain is geared to employ either approach depending on the context. For example, Daw et al. (2005), Lee et al. (2014), and Dunne et al. (2016).

indicator-based forecasting structure (as in He and Kučinskas (2024)[5]) and an emphasis on DGP extraction (as in Kendall and Oprea (2024), which we discuss below), making the convergence of findings across such diverse experimental designs particularly encouraging.

The study also ties into the broader literature on bounded rationality, which offers a diverse array of mechanisms to explain this cognitive duality. A first class of explanations involves rational inattention, where participants trade off accuracy against processing costs—so-called "*frictions*" (Handel and Schwartzstein, 2018). Specifically, Gabaix (2014)'s model suggests sparsity-based simplification whereby participants would choose to ignore the aberration feature—a mechanism supported by evidence linking 'sparsity complexity' to DGP extraction (Kendall and Oprea, 2024). Afrouz et al. (2023) provides additional insights through rational inattention to past data (the agent would only focus on the latest observation), though this mechanism likely played a limited role in our experiment by design—participants had free access to the asset value history. Alongside these friction-based accounts, the literature highlights several "*mental gaps*" (Handel and Schwartzstein, 2018)[6] potentially at work in our setting. Of particular relevance is Gagnon-Bartsch et al. (2018)'s insight into how inaccurate a priori mental models can prevent us from noticing evident regularities in datasets (what they refer to as "statistical gorillas") through constraining attention. In our setting, participants' prior models may exclude aberrations, channeling their attention away from statistical patterns that would reveal their presence. Additionally, their perception of shock persistence may anchor to levels higher than the true value, See, for example, Gabaix (2019), He and Kučinskas (2024), and Acharya et al. (2023).[7] Bottom-up attention mechanisms may further be at play. In this vein, Bordalo et al. (2023)'s theory of divided attention suggests that large forecast errors, being perceptually prominent (as evidenced by Findling et al. (2021)[8]), monopolize attention at the expense of attending to system parameters, consistent with the aforementioned 'system neglect' (Kremer et al., 2011). Graeber (2023)'s "*nuisance neglect*"—impaired noise filtering during inference—may further contribute.[9]

By requiring participants to extract the task's underlying DGP, our experiment also contributes to the growing literature on mental models in economics—focusing on how individuals interpret their environment under fundamental uncertainty about its structure. Such deep uncertainty has re-

[5] As in He and Kučinskas (2024)'s experiment, our participants must use an indicator variable to predict a target variable, in contrast to prior univariate forecasting settings. Our setting differs, however, in using binary variables and requiring participants to learn an underlying pattern rather than a simple correlation.

[6] Following Handel and Schwartzstein (2018), we define frictions as chosen cognitive simplifications and mental gaps as involuntary limitations.

[7] Interestingly, Acharya et al. (2023) shows that, under deep uncertainty about shock duration outside the lab, such biases can constitute optimal responses rather than mental gaps.

[8] Findling et al. (2021) provides evidence for the "Weber-imprecision model" in which the agent's perceptual system magnifies large prediction errors.

[9] Extrapolative heuristics (Fan et al., 2021) and hardwired extrapolation (Frydman and Nave, 12017) may further drive overreaction to the aberration events in our setting, though these likely played secondary roles in our experiment given that we focused participants on forecasting asset value switches rather than values themselves.

ceived considerable theoretical attention, with studies showing how misspecified mental models can generate seemingly biased behavior and asset pricing anomalies (e.g., Barberis et al. (1998), Brav and Heaton (2002)), exploring the asset pricing implications of mental model switching (Hong et al., 2007), highlighting the inherent complexity of pattern detection (Aragones et al., 2005), and describing how decision-makers can become entrenched in misspecified models (e.g., Gagnon-Bartsch et al. (2018) and Ba (2021)). In contrast, experimental work has largely overlooked how people form mental models under deep uncertainty. Most experiments restrict themselves to parameter learning under known DGPs, while the few featuring deep uncertainty assume that participants either know the DGP or hold a fixed misspecified model.[10] A notable exception is Kendall and Oprea (2024), who pioneered the experimental study of DGP extraction by having participants uncover deterministic patterns in two-variable datasets. They find that only 20% of their participants managed to extract their 'Switch' process, which appears at odds with our finding that a majority of participants extracted our DGP, despite it being more computationally complex than their Switch process (as detailed in Section 2.1.1). In Section 4, we propose several hypotheses to explain this apparent contrast. At the same time, our results echo Kendall and Oprea (2024)'s finding that decision-makers favor simple models: a non-negligible share (11%) of pattern-aware participants behaved *as if* they ignored the aberration feature. This apparent simplification may have been instrumental in their strong task performance.

Participants not only extracted the DGP but also adopted the optimal (model-based) approach to estimate its key parameters—thus clearing both aforementioned hurdles. This finding aligns with the growing "active thinker" perspective in psychology, which views individuals as agents who actively strive to understand their environment for better decision-making (e.g., Szollosi and Newell (2020) and Szollosi et al. (2023)). This view stands in sharp contrast to the prevailing "cognitive miser" view, which depicts people as passive recipients of information (see, e.g., Kahneman and Egan (2011) and Evans and Stanovich (2013)).

Our finding that an information nudge enhanced participant performance in the second experimental treatment contributes to the ongoing debate on nudge effectiveness in behavioral economics. For example, Hallsworth (2022), Simmons et al. (2022), and Chater and Loewenstein (2023). The prevalent view has held that behavior can only be influenced via interventions targeting the fast, automatic "System 1" while bypassing the deliberative "System 2" (World Bank, 2014). In contrast, the recent active thinker view argues that suboptimal behavior stems from cognitive and motivational misalignments, correctable through high-agency interventions (Szollosi et al., 2023). In this vein, Mertens et al. (2022) suggest that information nudges, which require motivated agents, are more vulnerable to differences in individual goals across a population than System 1 interventions,

[10] For example, Asparouhova et al. (2009) and Payzan-LeNestour and Bossaerts (2015).

potentially reducing their overall effectiveness. These recent perspectives informed our decision to motivate participants through significant incentives (Section 2.1) and our hypothesis that the nudge in our second experimental treatment would help participants extract the task structure (Section 2.3). Our experimental findings lend credence to these recent theoretical propositions, demonstrating the potential efficacy of information nudges when coupled with appropriate motivation.

Moving from behavioral economics to the emerging 'machine psychology' field, which examines AI decision-making processes in parallel with human cognition,[11] our findings offer potential insights into human-AI collaboration. As AI continues to advance, research in this field is expanding rapidly, prompting greater awareness of the associated risks (Jones, 2023) and sparking a reevaluation of the roles humans will play. While concerns about AI replacing humans have motivated substantial literature around the "Man versus Machine" theme,[12] Cao et al. (2024) propose that AI capabilities and human cognition are complementary in generating financial forecasts. They envision a future of "Man + Machine" rather than human displacement in financial decision-making. Their research indicates that while AI forecasts often outperform those of human analysts, humans retain an advantage in uncertain or unfamiliar scenarios.

Our study extends Cao et al. (2024)'s insights, highlighting a potential human advantage in scenarios where agents can mitigate deep uncertainty through pattern recognition. This is noteworthy given that pattern recognition in complex, ambiguous environments is often considered a weakness of AI; see, e.g., Hasanhodzic et al. (2019) and Larson (2021). In contrast, we observe surprising pattern recognition abilities in our task participants, pointing to a potential human edge in such contexts. Furthermore, we present preliminary evidence that this advantage might persist even if AI can extract the task DGP (Section 3.3). We compare AI and human forecaster earnings in task simulations, modelling machine learning as model-free (based on recent computer science and 'machine psychology' research[13]) and human learning as model-based (reflecting our finding that the vast majority of successful participants appeared to use a model-based approach to the task). We find that simulated humans significantly outperformed the top-performing model-free algorithm, highlighting the potential advantage of human-like, model-based learning in complex environments. However, this advantage is highly fragile: introducing doubt (in the form of Enke and Graeber (2023)'s "*cognitive uncertainty*") into our simulated humans was sufficient to erase their performance edge. While suggestive, these findings represent an initial exploration of human-AI complementarity—combining sophisticated (model-based) human cognition with AI's 'noise-free' execution (discussed further in Section 4).

[11] For example, Horton (2023), Aher et al. (2023), and Yax et al. (2024).
[12] For example, Acemoglu and Restrepo (2018) and Webb (2019).
[13] For example, Gawdat (2021), Hayes et al. (2024a), and Hayes et al. (2024b).

2 Forecasting shifts under deep uncertainty in the laboratory

2.1 Experimental design

2.1.1 Forecasting task

Task interface and payoff rule. We developed a simple web-based video game (a demo of the game is available here) to study shift forecasting under deep uncertainty. The game presents a chart displaying an asset's value unfolding in real-time, with a new value realized at each trial (Fig. 1).

At the beginning of each trial, participants are asked to forecast whether the incoming asset value will switch. They must make their forecast within 10 seconds, with an omnipresent counter showing the remaining time.[14] The task is self-paced, allowing decisions at any point within the time limit. Immediately after their choice, participants observe the outcome and the new asset value display (Fig. 2A).

The reward structure (Fig. 2D) is designed to incentivize accurate forecasting. A correct switch forecast leads to winning \$3, while missing a switch results in losing \$3. For no-switch forecasts, a correct prediction wins \$1, and an incorrect one loses \$1. Participants can also choose to skip a trial by pressing a "Skip" button, resulting in no outcome for that trial.

Participants complete 200 trials. At the end, 40% of these trials are randomly selected, and the cumulative outcomes determine the final payment. This payment ranges from a minimum of \$10 to a maximum of \$150, in compliance with our lab's ethics guidelines (detailed in Section 3.1).

Stochastic structure of the task. During each trial's decision phase, participants can click within the vacant area beneath the asset chart. This action instantly reveals an indicator chart on screen (Fig. 2A, top graph). This indicator is crucial for forecasting the next asset value because the asset and indicator charts are linked through a synthetic process based on the following rules. Starting from a baseline value (0), the indicator either:

(a) Remains at baseline (probability 0.75): In this state, a shift cannot occur, but an aberration—defined as a switch immediately followed by a reversal, with reversal probability of 1—may occur with a small probability ($p_a = 0.08$).

(b) Departs from baseline to a positive value (departure probability $p_d = 0.25$): In this state, an aberration cannot occur, but a shift becomes likely. We refer to this as "the dangerous zone". Here, shift occurrence follows a Poisson process with an intensity of 1, resulting in a high shift probability ($p_s \sim 0.7$) for each trial in the dangerous zone. The indicator continues

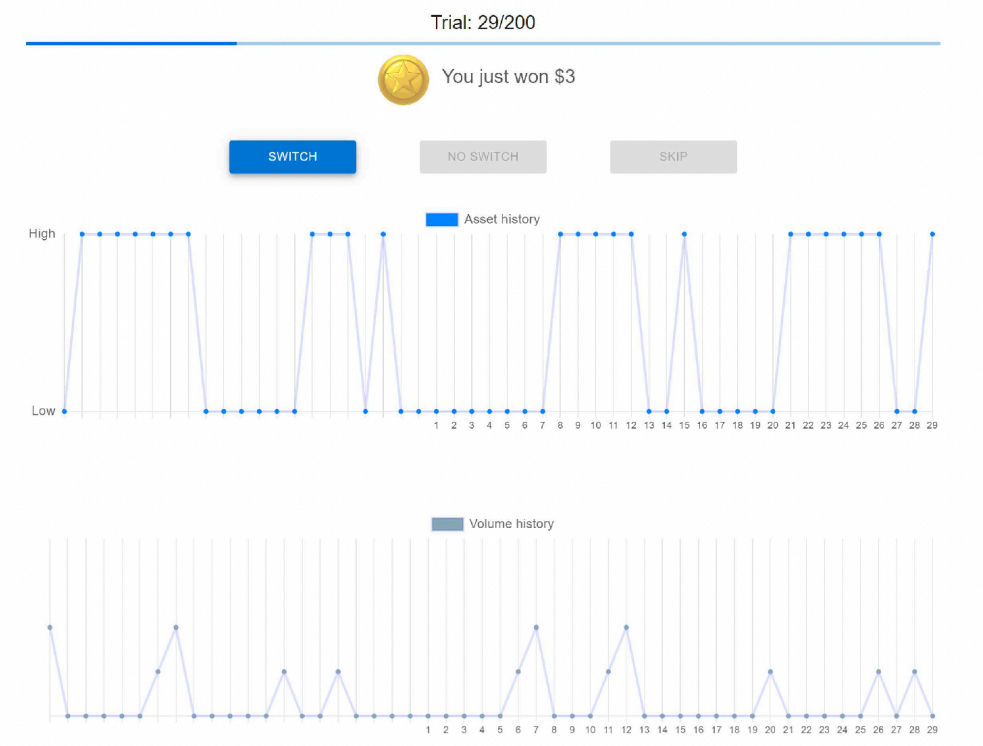
[14] Not answering within the 10-second time limit results in a \$1 loss. The task automatically terminates after 10 missed trials. In practice, missed trials were infrequent (see Table S3).

A

Choice (10 s)



Outcome (2 s)



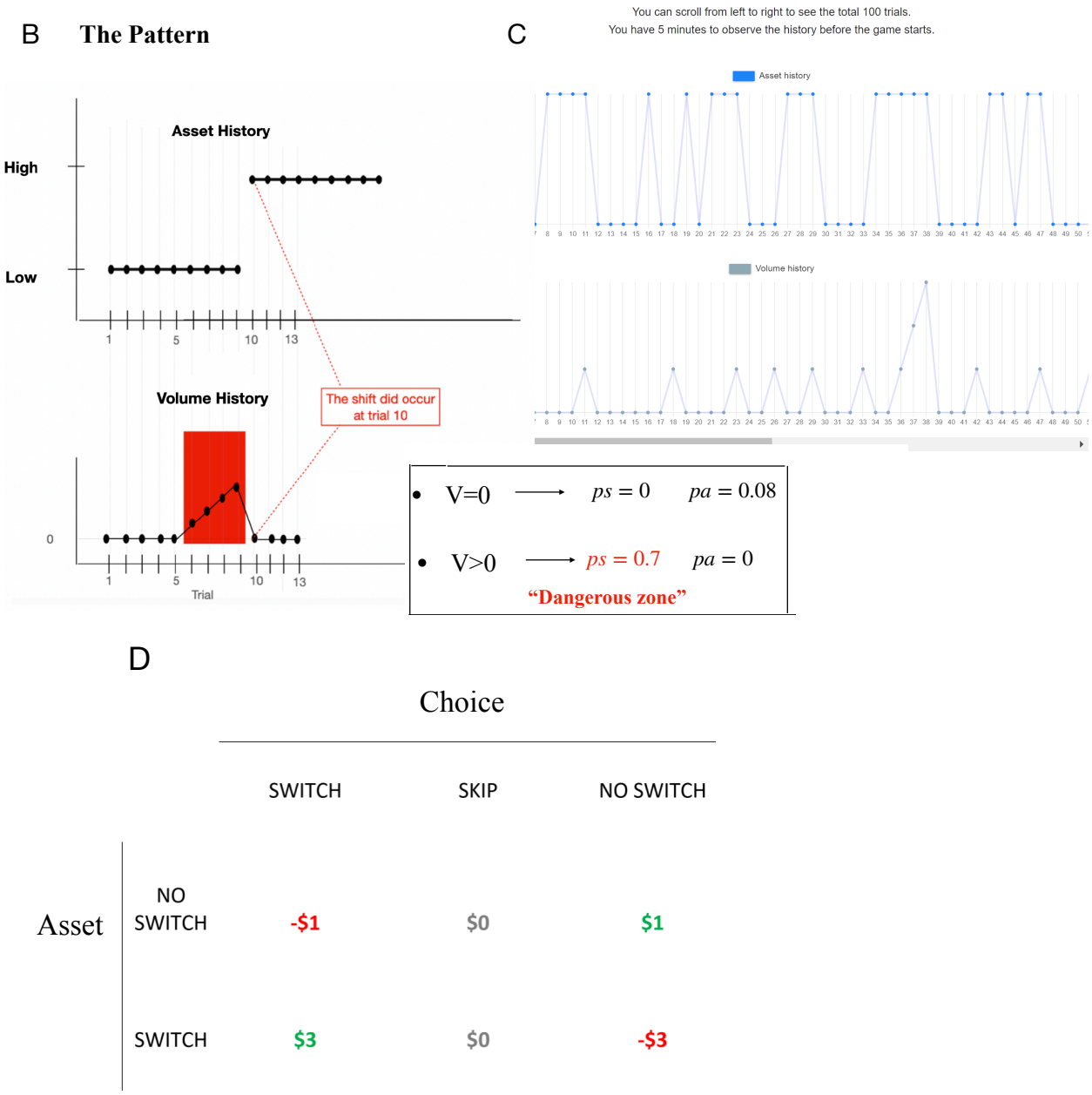


Figure 2. Experimental design. **(A)** Each trial of the task comprises a choice phase (10 s, self-paced) for asset switch forecasting, followed by explicit outcome feedback (2 s). During the choice phase, participants can reveal an indicator chart by clicking beneath the asset chart. **(B)** The indicator chart helps forecast shifts and differentiate them from aberrations (switches with immediate reversal, probability 1) for those who recognize the underlying pattern linking asset and indicator. Baseline indicator ($V = 0$) precludes shifts but allows rare aberrations (probability 0.08). Positive indicator ($V > 0$) signals imminent shifts (probability ~ 0.7)—we refer to these periods as "the dangerous zone." The task instructions provide no information whatsoever about the task structure to the participants. **(C)** Pregame, participants observe a 100-trial history of asset and indicator values, potentially uncovering the underlying DGP. The game starts immediately thereafter. **(D)** Payoffs are magnified for switch events: correct (incorrect) switch forecasts win (lose) \$3, versus \$1 in normal times.

to increase until a shift occurs, at which point it returns to baseline, ending the dangerous zone.

Fig. 2B provides a visual summary of this pattern linking indicator and asset values.

Extracting the task structure under deep uncertainty. Critically, the task instructions do not disclose the existence of this pattern to participants. This omission creates a complex learning challenge. In Section 2.2.1, we show that optimal play indeed requires learning both the shift probability (p_s) and aberration probability (p_a) within each relevant task context, as well as discerning the underlying stochastic process to identify these contexts. In essence, participants face deep or Knightian uncertainty, confronting an unknown structure in the task environment.

To provide participants an opportunity to mitigate this deep uncertainty through DGP-inference, we present them with a 100-trial history immediately before they begin the task. This pregame history displays the joint realization of asset and indicator values at each of 100 trials, randomly generated from the aforementioned synthetic process (Fig. 2C). Participants are allotted 5 minutes to observe this history before the game commences.

2.1.2 Information provided to the participant

First experimental treatment (T1). In the first experimental treatment (henceforth "T1"), participants faced full Knightian uncertainty, as the task instructions provided no information about the task structure. To justify the presence of the indicator on the screen during the pregame history phase (Fig. 2C), we employed a 'realism' rationale, alluding to the presence of similar indicators in real-world financial interfaces:

> "At the beginning of the game, you'll see what's happened with the asset in the last 100 days. You will also be shown the history for a volume indicator over the same period. This kind of indicator is commonly present on trading screens in the real world. Your job is to forecast in real time on each of 200 trials whether a switch is going to occur over the next day."

Second experimental treatment (T2). To address our second research question—whether DGP-extraction can be enhanced via an information nudge—we leveraged the high degree of control afforded by the laboratory experimental approach. We replicated the conditions of the original experiment with a second group of participants, replacing the original realism rationale with the following information nudge:

> "At the beginning of the game, you'll see what's happened with the asset in the last 100 days. You will also be shown the history for a volume indicator over the same

period. We are showing you this because your job is to forecast in real time on each of 200 trials whether a switch is going to occur over the next day, and there is a pattern linking the volume indicator to the regime shifts. If you manage to identify what the pattern is, you may track the value of the volume indicator to forecast the regime shifts by differentiating them from aberrations, which can greatly improve your performance."

To isolate the effect of this nudge, we ensured that both groups underwent identical experimental procedures and faced the same flow of events throughout the task. We also matched participant population characteristics across the two groups (more in Section 3.1.2).

Third experimental treatment (T3). Building on our prediction of the effectiveness of T2's basic nudge (explained in Section 2.3), we designed a third treatment (T3) to test whether nudged participants performed as well as participants fully informed about the task's DGP. While T3 maintained all T2 conditions, it added one key element: explicit instructions detailing the DGP in a simple, intuitive manner, ensuring that participants, regardless of their knowledge in statistics, could understand the statistical pattern central to the task.[15] As part of the experimental procedure, we verified that all T3 participants fully comprehended these explanations before commencing the task (more in Section 3.1.1). The T3 group thus serves as a control, providing a benchmark to assess the effectiveness of our nudge by comparing T2 and T3 task performance.

2.1.3 Rationale behind the current design

We developed our design based on three key criteria: real-world relevance of the task, tractability and robustness of the analyses, and statistical power of our empirical tests.

Task interface. The task interface aims to strike a balance between experimental control and real-world relevance, isolating the ubiquitous challenge described in the Introduction, which is particularly salient in financial investing. Rather than replicate existing financial displays, we distill this cognitive challenge to its essence, abstracting away from the incidental features of existing financial environments. This controlled setting allows us to isolate mechanisms that would be obscured in real-world markets.

The 'click' feature—where participants decide whether to display the indicator chart during each trial's choice phase (Fig. 2A, "Choice")—facilitates inference in our analysis of participant behavior. We leverage this 'click' decision data as an initial filter to classify participants who potentially discovered the pattern. We can confidently assert that participants who did not consistently click throughout the task failed to identify the pattern, allowing for preliminary segregation in our

[15] The script of the task instructions provided to the participants is available here.

analysis between potentially pattern-aware and demonstrably pattern-unaware participants. This initial triage enhances the precision of our subsequent analyses.

To ensure that the click feature did not introduce experimental artifacts (e.g., enhancing attention to the indicator or suggesting its usefulness), we conducted T1 sessions without this feature. In this variant, the indicator chart was a permanent element of the task interface alongside the asset chart. We found no behavioral differences across the two variants in terms of participant accuracy, earnings, and extent of pattern recognition. Therefore, to maximize statistical power for our main tests, we pool the data obtained from both T1 variants.

Payoff asymmetry. Our decision to magnify the payoff precisely when shift events occur adheres to all three design criteria. First, it captures a fundamental principle of real-world decision-making: the importance of decisive action during structural shifts. Financial markets illustrate this well—regime changes demand immediate response, while normal periods see minimal activity.[16] Second, it makes optimal behavior robust to agent's risk preferences (as shown by Monte Carlo simulations in Appendix A.3.1), reducing the confounding effect of individual risk attitudes in our analysis of participant behavior. Third, and most importantly, it maximizes optimal forecasters' earnings advantage over alternative strategies (without this feature, the earnings gap shrinks substantially—see Fig. S12, p.41). This matters for both external validity (trying to mimic real-world stakes where optimal action pays off) and statistical power: to infer participants' learning styles from their forecasts, we need distinct learning styles to generate different forecasts. Our payoff structure amplifies these differences, aiding inference.

Choice of task parameters. Our choice of task parameters (p_a, p_d, and p_s) and stochastic structure was guided by the same criteria—tractability, ecological relevance, and power. We tested several parameter combinations through Monte Carlo simulations of the task; the current specification appeared to be optimal for our purpose. We modelled the dangerous zone's shift process as Poisson primarily for tractability, as other specifications would complicate the study without altering the task's core features.

2.2 Behavioral models

The mathematical description of the models can be found in Appendix A.

2.2.1 Optimal behavior

— Learning *about* the environment.

[16] For example, Lo and Repin (2002), Lo et al. (2005), and Fenton-O'Creevy et al. (2011). This reality extends beyond finance. For instance, in ecology, species survival depends on swift adaptation to environmental changes, and in geology, appropriate responses to seismic precursors are life-saving.

As explained earlier, discerning the different contexts through pattern recognition is a prerequisite for optimal play. For T1 participants, this implies extracting the DGP from scratch (deep uncertainty); for T2, identifying it; and for the control group (T3), correctly interpreting it from the task instructions (full information about the DGP).

This pattern recognition qualifies as complex in these sense of Kendall and Oprea (2024). To be more specific, we denote by V_t the volume in trial t and introduce an indicator S_t for the asset movement in trial t, with $S_t = 1$ indicating an asset switch in trial t. Pattern-aware agents are to classify trials into the following contexts based on $(S_t, V_t, S_{t-1}, V_{t-1})$:

1. $V_t > 0, S_t = 0, V_{t-1} \geq 0, S_{t-1} \geq 0$: *inside the dangerous zone* [17]
2. $V_t = 0, S_t = 1, V_{t-1} > 0, S_{t-1} = 0$: *upon-shift*
3. $V_t = 0, S_t = 1, V_{t-1} = 0, S_{t-1} = 0$: *upon-aberration*
4. $V_t = 0, S_t = 1, V_{t-1} = 0, S_{t-1} = 1$: *upon-reversal*
5. $V_t = 0, S_t = 0, V_{t-1} = 0, S_{t-1} = 1$: *post-reversal/shift*
6. $V_t = 0, S_t = 0, V_{t-1} = 0, S_{t-1} = 0$: *normal times*

To perform this classification, the agent must attend to 2 inputs (current and previous volume values) and 2 outputs (current and previous asset movements), totaling 4 elements. Using the "sparsity complexity" metric proposed by Kendall and Oprea (2024) [which measures the number of input/output elements in the data set that the decision-maker must monitor to solve the task], we conclude that our task has a minimum sparsity complexity of 4.[18] This is a lower bound: due to the stochastic nature of our DGP, the realized data series may contain clustered switch events, such as when an aberration occurs immediately after a regime shift (e.g., Fig. 2C, trial 20). In such cases, the agent needs to look further back at realized values of both asset and volume, increasing sparsity complexity to 6.[19] Of note, the vast majority (74%) of our lab participants encountered data series with such clustered switch events.

— Learning *in* the environment.

[17] Participants may further distinguish between first, second, and third-plus trials within the dangerous zone. While just identifying 'inside the zone' is optimal (since shift probability remains constant at ps ~ 0.7), debriefing feedback suggests some pattern-aware participants used finer distinctions. Our analysis accommodates both cases, and our conclusions hold whether or not assuming within-zone differentiation.

[18] Alongside sparsity complexity, Kendall and Oprea (2024) propose the related metric of "partition complexity," which they find to be equally relevant for measuring complexity in DGP extraction. In our task, this metric corresponds to the number of contexts the agent must identify to solve the task—six at a minimum (see next).

[19] The DGPs in Kendall and Oprea (2024) are deterministic, so their complexity is fixed, whereas our DGP's complexity varies with the realized data stream in the manner just described.

Solving the task requires inferring the shock probabilities p_a and p_s (the optimal model-based approach) in each of the 6 contexts. The ensuing behavior is shown in Fig. 3. Note how the agent always forecasts 'Switch' (S) in the dangerous zone. This reflects the fact that $p_s > 1/4$ in this context (so forecasting S has positive expected value), which the agent estimates from the 100-trial pregame history and the early trials of the task.[20] Upon the shift and thereafter, in the absence of an aberration the agent forecasts 'No Switch' (N), which reflects the fact that outside the zone, p_s is null and the probability of an aberration is small ($p_a < 1/4$), which again the agent estimates using the pregame history and early trials of the game. In the event of an aberration, the agent knows about the immediate reversal which happens with probability 1 and hence they rightly forecast S. Likewise, in the post-reversal trials the agent forecasts N, which reflects his learning that p_s is null and p_a is very small in that context.

2.2.2 Model-based participants who failed to recognize the pattern ("non-recognition model-based")

Without pattern recognition, participants may only consider post-switch and no-switch contexts, or no context at all. While this reduces the complexity of the task, it is prohibitive for performance. Indeed, when applied to incorrect contexts, the optimal probabilistic approach proves highly ineffective (simulation results are shown in Fig. S12A, p.41). To be more specific, model-based participants who fail to learn the relevant task contexts estimate switch probability in one of two ways: i) separately for 'a shift has just occurred' and 'no shift has just occurred' contexts, or ii) in a context-free manner. The latter variant more accurately describes participant behavior (see Appendix A.3.2 for details), and we therefore adopt it for our main analysis: the agent estimates switch probability by recording switch vs. no-switch instances (combining shifts and aberrations), using pregame history and early task trials. Their estimated switch probability thus reflects the unconditional switch probability in the task, which is just above 1/4 by design, leading them to always forecast S (Fig. 3).

2.2.3 Adaptive participants who failed to recognize the pattern ("non-recognition adaptive")

Given the poor performance of applying model-based learning under a misspecified DGP, those who failed to extract the DGP may have opted for a more straightforward adaptive approach (model-free Reinforcement Learning). Adaptive behavior can take many forms in the current task. One key distinction is whether the adaptive agent implements model-free Reinforcement Learning (RL)

[20] Note that we assume risk neutrality, no cognitive uncertainty, and no "trembling" (choice randomness) in this section. These assumptions are relaxed in our analysis of participant behavior.

on a context-by-context basis or in a context-free manner.

EWA. For somebody who failed to recognize the pattern, a context-free approach may seem natural, and we use the generic "Experience-Weighted Attraction" (EWA) framework to formalize this case (Camerer and Hua, 1999). EWA encompasses a spectrum of adaptive approaches, varying based on whether the agent learns from counterfactual outcomes in addition to experienced ones, and the extent of past discounting, among other learning aspects (See Appendix A.2 for details). We consider all of them in our investigation of participant behavior. Simulated EWA behavior consistently features the patterns shown in Fig. 3. The agent in effect implements the "win-stick lose-switch" heuristic (echoing Charness and Levin (2005)'s terminology): a given action tends to be reinforced if it recently yielded a positive outcome; in case of a negative outcome, the agent switches to a different action. This explains why the agent typically struggles—in the zone, upon a shift, and upon a reversal—with one exception: upon an aberration. In this case, because forecasting N yielded a negative outcome in the previous trial, the agent inadvertently makes the correct decision by forecasting S.

Non-recognition context-specific RL We also considered the possibility that participants who failed to grasp the task's relevant contexts applied RL separately to the two "default" contexts—'a shift has just occurred' and 'no shift has just occurred'—using the standard adaptive expectations rule (which has been widely adopted to model learning in various disciplines[21]). Like EWA, this model encompasses a wide range of variants, depending on whether the agent employs counterfactual learning, adjusts the learning rate, or incorporates other features (summarized in Fig. 8, Appendix A.1.1). We tested all possible variations within our adaptive expectations algorithm and selected the best-fitting variant for the main analysis (see Appendix A.1.1 for details).

The simulated behavior consistently exhibits the suboptimal patterns shown in Fig. 3. Notably, the agent erroneously forecasts N upon encountering an aberration (Fig. 3B, "upon aberration"). This occurs because the decision relies on the most recent switch outcome, whether a shift or reversal, where Forecast N was previously rewarded by design. Similarly, upon an aberration reversal, the agent incorrectly forecasts S (Fig. 3B, "upon reversal") because Forecast S has been rewarded in the directly preceding aberration trial.

[21]For example, Evans and Honkapohja (2001) (in economics) and Daw (2014) (in neuroscience).

2.2.4 Adaptive participants who managed to recognize the pattern ("pattern-aware adaptive")

Participants could apply the same adaptive expectations algorithm, but to the task's relevant contexts (those described in Section 2.2.1), thereby optimizing the "learning *about*" dimension but not the "learning *in*" dimension (using adaptive rather than probabilistic learning). As with the non-recognition case, we considered variations in outcome appraisal (learning from counterfactual outcomes or not) and learning rate specification for our adaptive exptectations algorithm, and selected the best-fitting variant for our main analysis (see Appendix A.1.1 for details).

This chosen variant closely mimics optimal behavior overall while presenting distinct features, apparent in Fig. 3. Note how the agent optimally forecasts N outside the zone absent an aberration. This reflects the fact that in the context 'outside the zone, no aberration', forecasting N previously yielded good outcomes (as no switch occurs in this context by design), increasing Forecast N's value. The agent also mimics optimal behavior upon an aberration (forecasting S) and subsequent reversal (forecasting N), due to identifying the relevant task contexts and learning that upon an aberration, Forecast S previously yielded a good outcome, while in the subsequent reversal, forecasting N did.

However, unlike the optimal agent, this type tends to forecast N in the initial trials of the dangerous zone (see Fig. 3A, "in zone"). This occurs when the previous zone lasted more than one trial and forecasting S in the first trial yielded a negative outcome, leading the agent to forecast N the next time this context is encountered. This anomaly occurs in 1/3 of the dangerous zones.[22] The second notable deviation from optimal behavior is the agent's forecasting S in the first trial outside the zone following an aberration (see Fig. 3B, "post reversal"). This systematic oddity reflects that the last encounter with this context was just before the aberration, where forecasting N led to a loss, thus decreasing its value. These departures from optimal behavior significantly impair earnings (simulation results are shown in Fig. S12A, p.41).

Summary

Our design and simulations highlight how, under deep uncertainty, extracting the DGP and learning its parameters are synergistic. Failing either dimension substantially impairs performance relative to the optimal benchmark. The analysis further shows distinct behavioral patterns across the five participant types, enabling us to reliably infer each participant's type based on their observed behavior in the lab.

[22] This frequency reflects the fact that about 1/3 of the zones last more than one trial by design (increasing the frequency of long zones increases the anomaly's prevalence).

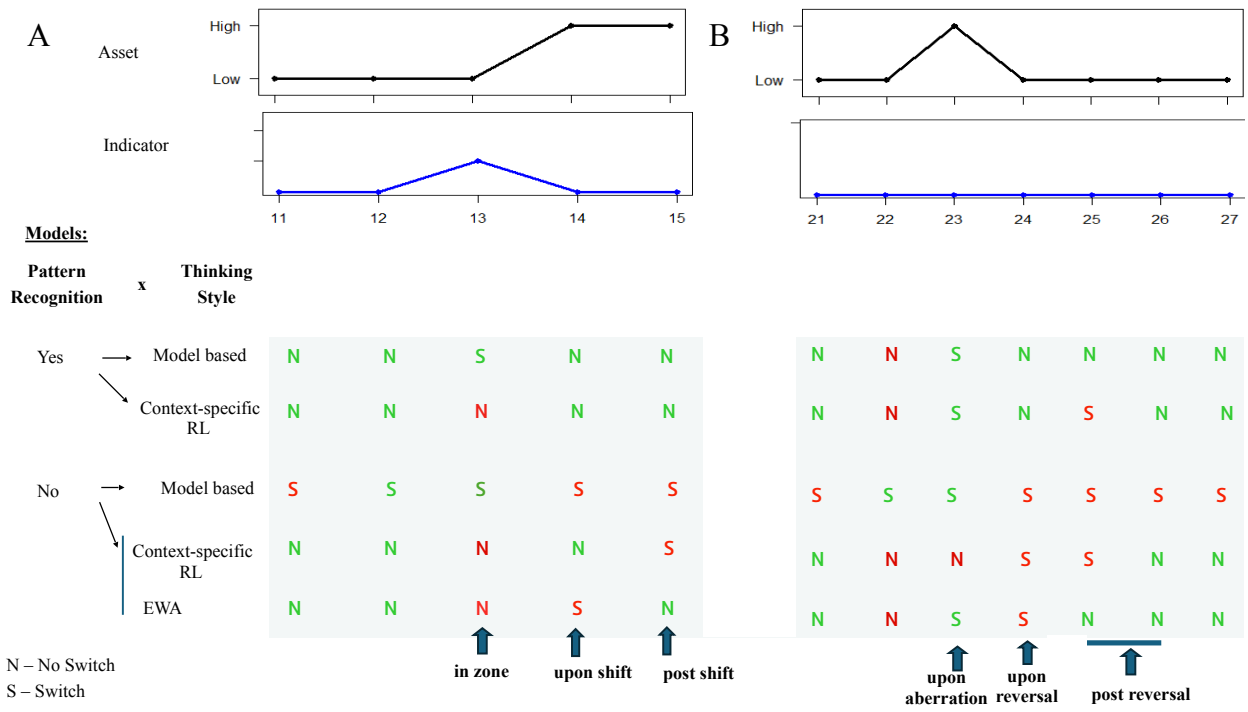


Figure 3. Model signatures. The figure illustrates how the five participant types behave during a shift (**A**) and an aberration (**B**).

2.3 Preregistered hypotheses and tests

The hypotheses, methods, and analysis plan were preregistered on the Open Science Framework (OSF) on 13th June 2023, prior to data collection in the lab which began on 14th June 2023.[23]

2.3.1 Main hypotheses

H1. In T1, the average participant will fail to detect the pattern, resulting in suboptimal behavior.

H2. In T2, the average participant will detect the pattern, leading to near-optimal behavior.

H3. In T3, the average participant will correctly interpret and apply the pattern explained in the task instructions, resulting in near-optimal behavior.

H4. Participant behavior in T2 will be superior to that in T1.

H5. Participant behavior will be similar in T2 and T3.

Rationale. These hypotheses align with the recent "active thinker" paradigm in psychology, which posits that humans are proactive agents who actively seek to understand their environment and can learn from provided information (as discussed in the introduction). Influenced by this

[23]The OSF preregistered file for this study is available here.

paradigm, we began the study with a strong prior belief that:

(i) The control group (T3) would effectively learn from the provided information (**H3**).

(ii) The nudged group (T2) would independently discover this information (**H2**).

(iii) Consequently, T2 and T3 groups would exhibit similar behavior (**H5**).

We further hypothesized that T1 participants would typically fail to independently discover the pattern (**H1**). Notably, the sparsity complexity of our task (minimum of 4; see Section 2.2.1) exceeds that of Kendall and Oprea (2024)'s "*S2*" DGP (complexity of 2). Moreover, the stochastic nature of our task introduces an additional layer of difficulty (Oprea, 2024). Given that only 22% of participants in Kendall and Oprea (2024) successfully identified *S2*, it is reasonable to expect a higher failure rate in our more demanding task, thus supporting the plausibility of **H1**.

The combination of **H1** and **H2** naturally led to **H4**, predicting superior performance in T2 compared to T1.

2.3.2 Statistical tests

Estimation procedure. To test our hypotheses, we determined each participant's profile among the five types described in Section 2.2, by fitting each of our five behavioral models to participant behavior and comparing their goodness of fit.

Parameter fitting. For each model, we generated trial-by-trial action values, calculated action probabilities based on these values using the logit rule, and maximized the log likelihood (LL) of the participant's observed choices for the first 100 trials of the game (training dataset).[24]

Model Comparison ("horse race"). We calculated each fitted model's LL for the second 100 choices (out-of-sample test). This out-of-sample LL served as our primary measure of model goodness of fit. We also computed the model's out-of-sample accuracy (percentage of correct model predictions of participant's observed choices, averaged over 10 repetitions) as a complementary metric.

We opted for this two-step procedure over simpler approaches (which correct for model complexity through standard criteria such as BIC) because out-of-sample testing is believed to more effectively penalize model overfitting in model comparisons of this nature; see, e.g., Ho et al. (2007, 2008). Crucially, our procedure demonstrates excellent "*recovery*" properties in the sense of Palminteri et al. (2007), as shown in Online Appendix B. For instance, we simulated 72 optimal

[24]The optimal model, pattern-aware RL model, non-recognition RL model, and non-recognition model-based model each have 3 free parameters, while EWA has 8; see Appendix A for details.

players and found that the optimal model provided the best fit for over 95% of them. Conversely, for 72 simulated non-recognition participants, the optimal model almost always underperformed one of our three non-recognition models, demonstrating satisfactory "*restrictiveness*" in the sense of Fudenberg et al. (2023).

Main statistical tests.

Test 1 (N=71): We conducted a paired t-test comparing the goodness of fit of the optimal model versus the best-fitting non-recognition model for each T1 participant. The null hypothesis (H0) posits that the goodness of fit is superior or equal for the optimal model. We anticipated rejecting H0, consistent with our core hypothesis **H1**.

Test 2 (N=62): We applied the same paired t-test to the T2 group of participants. H0 states that the goodness of fit is superior or equal for the best-fitting non-recognition model. We anticipated rejecting H0, consistent with **H2**.

Test 3 (N=133): We employed a χ^2 test for proportions to compare the prevalence of pattern recognition in T2 versus T1. H0 posits that prevalence is inferior or equal in T2. We defined pattern recognition as prevailing for a participant if the optimal model fits their behavior better than the best-fitting non-recognition model. We anticipated rejecting H0, consistent with **H3**.

Test 4 (N=60): We conducted the aforementioned paired t-test on the T3 group of participants. H0 states that the goodness of fit is superior or equal for the best-fitting non-recognition model. We expected to reject H0, in accordance with **H4**.

Test 5 (N=122): We used the same approach as in Test 3, this time comparing the prevalence of pattern recognition in T2 versus T3. H0 posits that the percentage in T2 equals the percentage in T3. We anticipated being unable to reject H0, consistent with **H5**.

Importantly, as part of the preregistration plan for the study, we verified that each test has satisfactory a-priori power ($> 90\%$ chance to detect a medium effect size).

In addition to these statistical tests, we further quantified the relative evidence for each model. For each participant, we computed the posterior probability that they behaved according to each model, inspired by Dal Bó and Fréchette (2011):[25]

$$\hat{p}_n(k) \propto e^{L_n(k)} \cdot p_{n,k}, \qquad \sum_{k=1}^{4} \hat{p}_n(k) = 1, \tag{1}$$

where k indexes the four models—the optimal model, non-recognition RL model, non-recognition

[25] Dal Bó and Fréchette (2011)'s "*Strategy Frequency Estimation Method*" is designed for games with deterministic strategies, binary actions, and a single free parameter per task, so we adjusted it to our setting. We thank an anonymous reviewer for suggesting this additional analysis.

model-based model, and EWA model.[26] $p_{n,k}$ denotes the prior probability that participant n adopted model k, and $L_n(k)$ denotes the out-of-sample LL of model k for participant n. We used an agnostic prior (optimal model: $\frac{1}{2}$; non-recognition: $\frac{1}{2}$, $\frac{1}{6}$ for each of the 3 non-recognition models) in the main analysis, and checked the conclusions are robust to using different priors.

3 Results

3.1 Experimental procedures

The experimental protocol was approved by the University of New South Wales Human Research Ethics Committee (project number HC230306). Informed consent was obtained from all participants, who were undergraduate students (N=193) from the University of New South Wales. Prior to the session, they were instructed to study online task instructions [available here]. These instructions introduced the video forecasting game, delineated its rules, and mentioned the pregame history phase, stating: "*At the beginning of the game, you'll see what's happened with the asset in the last 100 days.*"

3.1.1 Incentivization Scheme and Transparency of Task Instructions

Crucially, the instructions emphasized our experimental design's central incentivization scheme (explained in Section 2.1.3, "Payoff asymmetry"):

> " You can earn a significant amount of money from the task (up to \$150 AUD[27]) if you perform well, but payments will also be varying considerably across participants. To be more specific, if you play optimally, you will be mostly likely to leave the lab with more than \$130 AUD, and even highly likely to get the max payment of \$150 AUD. However, if you do not implement the optimal strategy for the task, your most likely payoff from the task will be \$5 AUD."

This aspect of the task instructions seeks to replicate real-world incentives in the lab as closely as possible. It complements our strategy of designing a task where striving for optimal performance is objectively worthwhile, yielding significantly better outcomes (see Fig. S12A, p.41). We considered

[26] Of note, we exclude the pattern-aware adaptive type described in Section 2.2.4 from our main tests to adhere to our preregistration plan. Including it, however, further strengthens our main findings (details follow in Section 3.2.2).

[27] The \$150 cap applied to the first experimental treatment (T1), while T2 and T3 had a \$100 cap. Although we initially planned a uniform \$100 cap, the UNSW Human Research Ethics Committee required a higher cap for T1 "in the interest of fairness to the T1 participants, as the task is effectively infeasible for them." To preserve strong participant motivation in T2 and T3, we opted to raise the T1 cap rather than lower the others. Given that \$100 is already a significant amount for most students, we expect this differential to have minimal impact on motivation across treatments.

it crucial to inform participants of this distinctive feature to ensure full engagement, particularly given the widely held view in experimental economics that incentivizing participants is vital for complex cognitive tasks.[28]

Upon arrival at the laboratory, participants reviewed the online instructions for 20 minutes and completed a multiple-choice questionnaire to verify their comprehension. The experimenter then discussed the answers with them for 10 minutes. Following this, participants received a final briefing on the payment procedure and central incentivization scheme, reinforcing the written instructions. T3 participants additionally received a verbal reminder about the task's stochastic structure. Finally, all participants completed one task run, averaging 30 minutes, including the 5-minute pregame history review.

3.1.2 Matching procedures.

To ensure an unbiased comparison of participant behavior across treatments, we implemented two procedures. First, we matched participants' key demographic characteristics (age, gender, academic background). Participants completed a brief questionnaire, selecting from predefined categories for academic major (business school, science, other), gender, and age. We maintained balanced proportions across treatments for each characteristic (Table S2).

Second, we ensured that participants across all three treatment groups experienced identical event chains. We generated 72 data series, each containing 300 trials (100 pregame history + 200 task trials), all derived from the task's stochastic structure. Each series was randomly assigned to one participant in each treatment group (T1, T2, T3). This approach balanced control and replicability concerns. (NB: Using a single data series for all participants was deemed unsuitable due to replication concerns.)

3.1.3 Precluding information spillover.

Given the nature of our study, controlling information available to participants outside the lab was a crucial aspect of our controlled experimentation. To prevent potential "contamination" between treatments (e.g., T3 participants revealing the task's nature to those in T1 or T2), we conducted the treatments sequentially: T1 first and in a single batch, followed by T2, and T3 last. Additionally, we framed each treatment as an independent experiment to prevent participants from perceiving any connection between them.

[28] For example, Wilcox (1993), Hertwig and Ortmann (2001), Charness et al. (2010), Payzan-LeNestour and Bossaerts (2015) and Oprea (2024). This perspective is also reflected in the "active thinker" paradigm discussed earlier.

3.2 Main findings

3.2.1 First-pass analysis

We first present the data using intuitive benchmarks, comparing the lab participants' performance against two reference points: on one hand, the ideal participant who behaves optimally with risk neutrality, no cognitive uncertainty, and no trembling-hand errors (as described in Section 2.2.1); on the other, participants who fail to extract the pattern and hence rely on simple heuristics (Gigerenzer and Todd, 1999; Fan et al., 2021) as a last resort. Among the latter, Heuristic 1 involves always forecasting 'No Switch'— mimicking (and, as it turns out, outperforming) our non-recognition RL model. Heuristic 2 involves always forecasting 'Switch', displaying behavior representative of our non-recognition model-based agents. Heuristic 4 consists of making purely random choices, while Heuristic 3, the worst performer, is the "win-stick lose-change" heuristic characteristic of the EWA model.

We simulated our lab experiment with each archetype using the 72 data series encountered by our lab participants. Table 1A summarizes their performance in terms of forecast accuracy during normal times and switch events. The results demonstrate that only the ideal agent performs well in both scenarios, while all heuristics fail to do so. Based on these findings, we conducted an initial analysis by classifying lab participants into two groups according to their cumulative outcomes, using \$60 as a threshold. Table 1B shows that participants in the 'Higher Performers' group (outcomes $\geq$ \$60) tend to perform well in both scenarios, mirroring the ideal archetype, while those in the 'Lower performers' group typically show poor performance in at least one scenario (Fig. S16), like pattern-unaware players.

The use of a \$60 threshold for this initial classification stemmed from several considerations. First, we expected optimal human players in the lab to achieve lower outcomes than the ideal archetype due to well-documented psychological factors, such as "*complexity-driven mistakes*" (Oprea, 2024; Vieider, 2024). For instance, they may experience cognitive uncertainty, which leads to belief compression, as modeled by Enke and Graeber (2023). To illustrate, consider an optimal agent who, due to cognitive uncertainty, anchors on 'No Switch' 30% of the time in the zone and upon an aberration (i.e., provides the correct forecast 'Switch' only 70% of the time) Our simulations show that such an agent's outcomes can drop to as low as \$60. Additional established factors, including loss aversion, which in our design leads agents to sometimes skip (counted as a forecast error), can further decrease performance.[29] As a result, cumulative outcomes around \$60 can reflect near-optimal behavior, where optimal cognition is "tainted" by these psychological factors. Conversely, the cumulative outcomes of someone who did not recognize the pattern are

[29] See Appendix A.3 for the version of the optimal model that allows for risk/loss aversion.

below \$60, regardless of their type (see Fig. S12A).

Table 1. Performance of simulated archetypes (Panel A) and the actual lab participants (Panel B). Numbers represent means. The 'Higher Performers' group consists of participants with cumulative outcomes of at least \$60.

		Forecast accuracy (%)		Cumulative outcomes (\$)
		Normal times	Switch events	
Panel A: Simulated participants:				
	Ideal Player	100	100	167
Pattern-unaware Player	Heuristic 1: *always no switch*	100	0	10
	Heuristic 2: *always switch*	0	100	-10
	Heuristic 3: *win-stick, lose-change*	68	20	-22
	Heuristic 4: *random*	67	34	-2.9
Panel B: Lab participants:				
Treatment 1	**Higher performers**	89	77	125
	Lower performers	63	49	18
Treatment 2	**Higher performers**	93	77	143
	Lower performers	79	37	28
Treatment 3	**Higher performers**	95	84	164
	Lower performers	52	65	8

The proportion of Higher Performers, as just defined, reaches 63% in T1, 89% in T2, and 93% in T3 (Fig. 4). Visual inspection highlights strong performance already in T1, thereby falsifying H1. Notably, the Higher Performers group in T1 achieves mean accuracy for switch events and cumulative outcomes of 77% and \$125, respectively, comparable to an ideal player bounded only by cognitive uncertainty as modeled above (70% and \$130). Fig. 4 also indicates a marked improvement from T1 to T2, with similar performance levels observed between T2 and T3.[30] Importantly, these comparative results remain robust to alternative threshold values used to distinguish Higher and Lower Performers.

3.2.2 Falsification of Hypothesis 1

Test 1 results confirm the falsification of H1. We cannot reject the null hypothesis ($t = 3.38$; $p = 0.9994$, one-tailed; Cohen's D = 0.4; Fig. 5A). In fact, there is evidence that the optimal

[30] See Table S3 for additional summary statistics, pooling all participants by treatment.

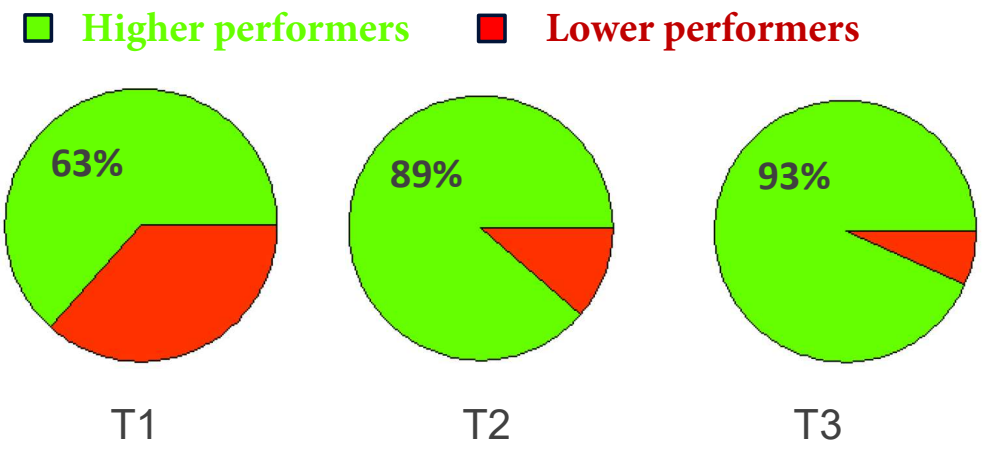


Figure 4. Participant classification based on cumulative outcomes. Pie chart showing the proportion of participants classified as 'higher performers' (cumulative outcomes $\geq$ \$60) in each treatment.

model provides a superior fit (same test, null hypothesis is that the optimal model's fit is inferior to that of the best competitor, $p < 0.001$). Notably, the optimal model fits better for 62% of participants, who were accordingly classified as pattern aware.[31]

As a validation check, we examined the click frequency data. While the majority of T1 participants clicked close to 100% of the time, 20% never clicked (Fig. S15). Reassuringly, these demonstrably pattern-unaware participants were also independently classified as such by our statistical tests.

To further investigate the learning process of these pattern-aware participants, we conducted a follow-up paired t-test on this subgroup, comparing the goodness-of-fit between the optimal model and the pattern-aware RL model. The results lead us to reject the null hypothesis of equal goodness-of-fit ($N = 44$, $t = 8.19$; $p < 0.001$, two-tailed; Cohen's D = 1.23). The optimal model provides a superior fit for a large majority (89%) of participants, suggesting that the pattern-aware group in T1 employed a model-based approach to the task.

We also examined the learning trajectory of pattern-aware participants: did they extract the DGP during the pregame phase (allowing them to behave near-optimally from the onset of the game), or did they gradually discover it over the course of the task? To gain preliminary insights into this question, we re-ran the model comparison for these participants using only the first 100 trials of the task. We found that 73% of them were classified as pattern-aware based on their behavior in this subsample, suggesting that they successfully extracted the DGP during the pregame phase, enabling them to act near-optimally from the start.

[31] This percentage underestimates the true prevalence of pattern recognition in T1, as our preregistered test conservatively excludes the pattern-aware RL model from the comparison between pattern-aware and non-recognition models. When we include the pattern-aware RL model in the 'horse race', the percentage increases to 68%.

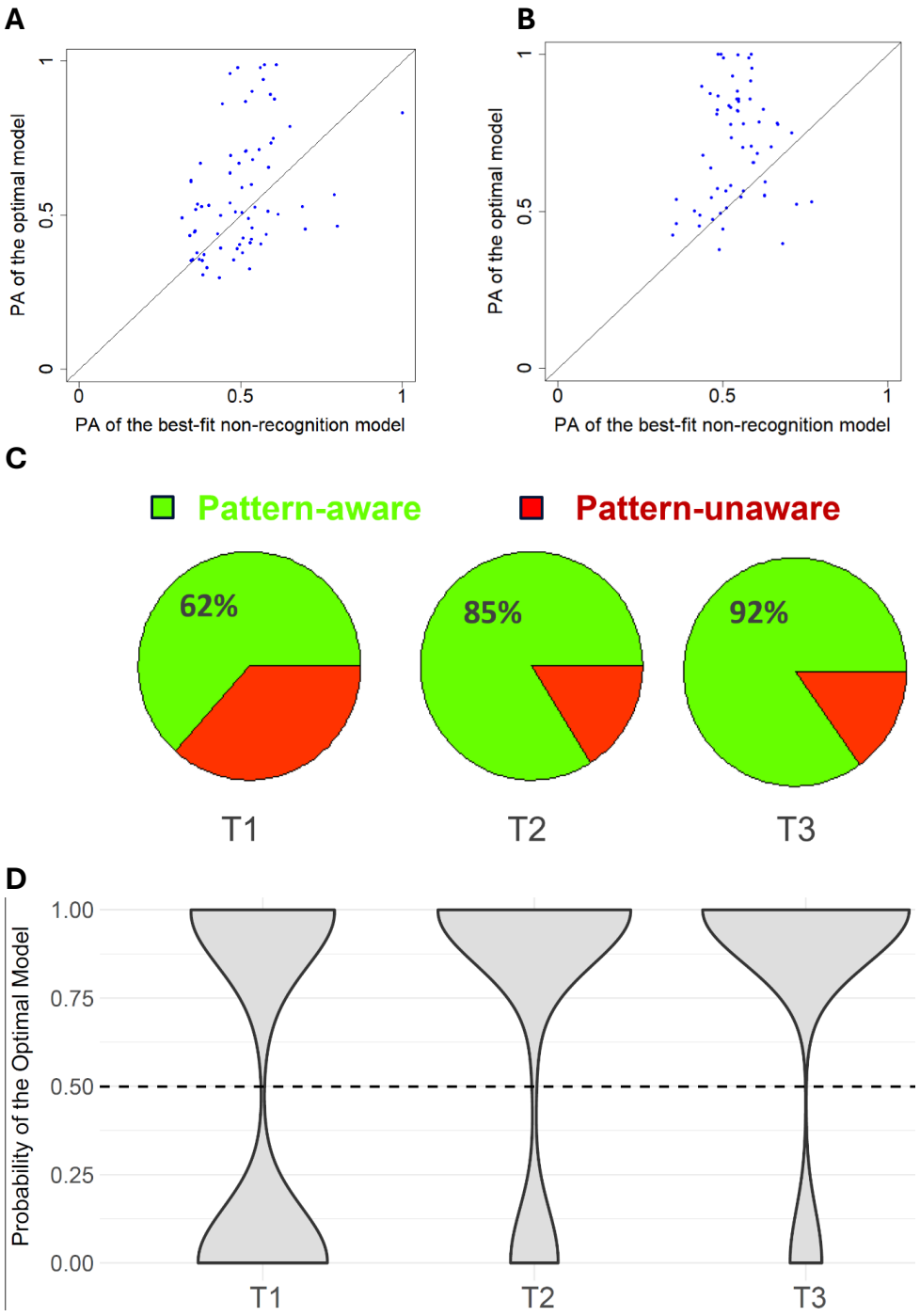


Figure 5. Comparison of model fit and pattern awareness across treatments. This scatterplot compares the prediction accuracy (PA) of the optimal model and the best-fit non-recognition model for participants in treatments T1 (**Panel A**) and T2 (**Panel B**). Points above the 45-degree line indicate a better fit for the optimal model. (**C**) Pie plot showing the proportion of pattern-aware participants (optimal model fits better) in each treatment. (**D**) Posterior probability of the optimal model for each participant, computed by Eq. (1). A probability above 0.5 indicates evidence favoring the optimal model, with values closer to 1 indicating stronger evidence. The violin plot's width represents the density of participants at each probability value.

3.2.3 Effect of nudging

We can reject the null hypothesis in **Test 2** ($t = 7.08$; $p < 0.001$, one-tailed; Cohen's D = 0.90; Fig. 5B). This finding suggests that the average T2 participant successfully recognized the pattern, with 85% of the T2 sample classified as pattern-aware based on the test results. Further analysis indicates that these participants employed a model-based approach, similar to their counterparts in T1. This conclusion was drawn from a follow-up paired t-test identical to the one conducted for the T1 group ($N = 53$, $t = 13.60$; $p < 0.001$, two-tailed; Cohen's D = 1.87).

Fig. 5C summarizes pattern awareness across treatments based on the above results. The plot closely mirrors the one based on the simple categorization derived from summary statistics (Fig. 4). Participant behavior in T2 appears superior to that in T1 (**Test 3** results: $\chi^2(1) = 8.12$, $p = 0.002$, one-tailed; Cramer's V = 0.25) and comparable to that of the control group (T3) which received full disclosure of the task structure (**Test 5** results: $\chi^2(1) = 0.62$, $p = 0.43$, one-tailed; Cramer's V = 0.07). These findings highlight the effectiveness of the information nudge implemented in T2.[32]

Summary of key findings. Participants not only learned from provided information (T3) and discovered information (T2), but also from information they independently sought out (T1), contrary to our initial expectations. Fig. 5D reinforces these conclusions by showing the posterior probability of the optimal model for all participants. Notably, 97% of participants exhibit a probability value either above 0.95 or below 0.05, offering clear-cut evidence for or against the optimal model.[33] The graph shows that the prevalence of the optimal model is higher in T2 than in T1 and does not increase further in T3 compared to T2.

Robustness checks. We verified the robustness of our findings by re-running all tests with alternative model comparison methods. Notably, for **Test 3** and **Test 5**, we improved inference reliability by combining accuracy and LL criteria, and applying a 'minimum difference' threshold for model comparisons. All methods corroborated and reinforced the results documented in Sections 3.2.2 and 3.2.3; see Online Appendix B for details.

3.2.4 Evidence of cognitive duality in participants

Further analysis of the pattern-aware participants' data reveals a clear asymmetry in forecasting performance: they excelled at predicting regime shifts but often failed to anticipate the self-correction following aberrations (Fig. 6A). As noted in the Introduction, this aligns with prior

[32] As expected, **Test 4** results indicate that the T3 group correctly interpreted and applied the pattern explained in the task instructions ($t = 10.91$; $p < 0.001$, one-tailed; Cohen's D = 1.41). According to the test results, 92% of the T3 sample demonstrated pattern awareness in their behavior.

[33] This classification criterion follows Kendall and Oprea (2024).

evidence showing that people struggle with transitory shocks but adapt well to persistent ones.

Two related explanations arise from the bounded rationality literature. First, participants may have deliberately ignored aberrations to simplify DGP extraction, in line with Kendall and Oprea (2024)'s finding that agents favor simpler mental models. Second, they may have entirely overlooked aberrations due to prior beliefs that ruled out their possibility, preventing them from recognizing this feature in the asset history (Gagnon-Bartsch et al., 2018). Under both explanations, T1 participants who mishandled aberration events should have done so *systematically*.

Yet our data suggest otherwise: only 11% of the pattern-aware participants exhibited systematic errors upon aberrations (see Supplementary Table S4 in Online Appendix D), suggesting that additional mechanisms were at play. Consistent with this hypothesis, the observed forecasting asymmetry persisted—albeit in a weakened form—even when participants were fully informed of the DGP in T3 (Fig. 6B). What could explain these purported "complexity-driven mistakes" (Oprea, 2024)? Bordalo et al. (2023)'s theory of divided attention, combined with Kremer et al. (2011)'s evidence of "system neglect," points to attentional lapses: the surprising nature of aberrations (by design, given the small p_A) may have caused participants to momentarily overlook their transitory nature. This attentional failure may have been further exacerbated by "nuisance neglect"—the tendency to disregard nuisance features (Graeber, 2023). Building on Graeber (2023)'s finding that attentional nudges enhance participants' integration of such features, we conjecture that periodic alerts emphasizing the nature of aberrations could have improved T3 participants' performance.

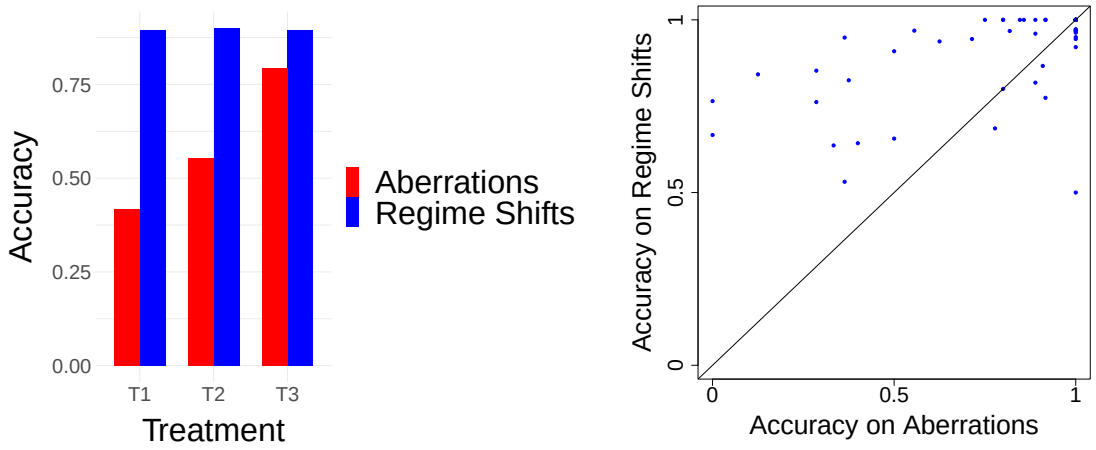


Figure 6. Cognitive duality in handling shocks. (**A**) Mean forecast accuracy by shock type for pattern-aware participants across treatments. (**B**) Individual forecast accuracy for pattern-aware participants in T3, comparing performance on regime shifts versus aberrations. Each point represents one participant; points above the 45-degree line indicate better performance on regime shifts than aberrations.

3.3 Opening the investigation of machine versus human forecasting ability

Our findings of pattern recognition in the T1 group contrast with AI's documented struggles in this area across various domains (Larson, 2021), suggesting a potential human advantage in T1-like environments where pattern recognition is critical.

However, the claim that humans outperform AI in such contexts warrants further empirical testing (see Section 4). Here, we consider an alternative scenario: If the machine recognizes the pattern, how does its performance compare to that of a sophisticated human forecaster?

To address this question, we compared the performance of pattern-aware machines and human forecasters in 500 Monte Carlo simulations of the task. Since pattern-aware participants consistently used a model-based forecasting approach across all treatments, we represented human forecasters using the optimal model. In contrast, AI is typically assumed to be model-free,[34] so we represented AI forecasting using an optimized RL algorithm tailored to the task. This algorithm's parameters and features were tuned to maximize task earnings (see Appendix A.1.2 for details) and served as the machine's representation in the contest against human forecasters. For both human and machine, we assumed risk neutrality and set the logit parameter to 5.[35] We verified the results remain qualitatively unchanged across varying levels of risk/loss aversion for the human forecasters and a meaningful range of logit parameter values.

Results reveal a significant human outperformance (paired t-test comparing mean cumulative outcomes: $N = 500$, $t = 27.18$, $p < 0.001$, 2-tailed; Cohen's D= 1.2, Fig. 7A). This finding was far from a foregone conclusion given the sophistication of our machine learning model. The machine was equipped with both observational and counterfactual learning capabilities—features associated with model-based learners (Dunne et al., 2016). In many contexts, incorporating these features into standard RL makes it equivalent to optimal (model-based) learning (Camerer and Hua, 1999; Pouget, 2007). However, in the current experimental context, this equivalence breaks down: despite enjoying the considerable advantage of unlimited pre-task access to task-generated data and fine-tuning of key learning rate parameters (see Appendix A.1.2 for details), the machine was significantly outperformed by humans. This superior human performance arises from the machine's excessive path dependence and instability, which remain key limitations of even the most advanced adaptive approaches in dynamic environments.

To assess whether this human outperformance generalizes beyond our specific task, we conducted simulations across various configurations, systematically varying key parameters—probability of entering the zone (p_a), shift probability in the zone (p_s), and aberration probability (p_d). First, we varied p_a and p_s while fixing p_d at 0.25 as in the experiment. Fig. 7B shows that the human

[34] For example, Gawdat (2021), Hayes et al. (2024a,b).

[35] For reference, the fitted logit parameter value across pattern-aware participants in T3 is 5.4.

advantage remains robust across all scenarios, with effect sizes exceeding 0.5 as long as there is uncertainty about switch occurrence. This robustness arises because optimal forecasters are consistently rewarded for correctly forecasting S in the zone. In contrast, the machine, constrained by its adaptive nature, forecasts N in the first trial of the dangerous zone when the previously encountered zone exceeded 1 trial—a hallmark of the context-specific RL approach (Section 2.2). This underscores the limitations of purely adaptive strategies in dynamic environments.[36]

Crucially, we demonstrate that this human advantage over the machine learning algorithm is highly sensitive to cognitive limitations: incorporating realistic human "noise" into our simulated human forecaster eliminates its performance edge. Specifically, we introduced a moderate level of cognitive uncertainty into the human forecaster (following Enke and Graeber (2023)'s model), where the simulated agent incorrectly anchors on 'No Switch' 10% of the time in the zone and upon an aberration. This modification, reflecting documented human behavioral patterns, reverses our earlier findings: the machine learning algorithm now significantly outperforms the human forecaster (same paired t-test as before: $N = 500$, $t = -8.85$, $p < 0.001$, two-tailed; Cohen's D= -0.4).

[36] In complementary analyses, we varied p_d and p_s while fixing p_a at 8% as in the experiment. Results again show a consistent human advantage (see Fig. S14).

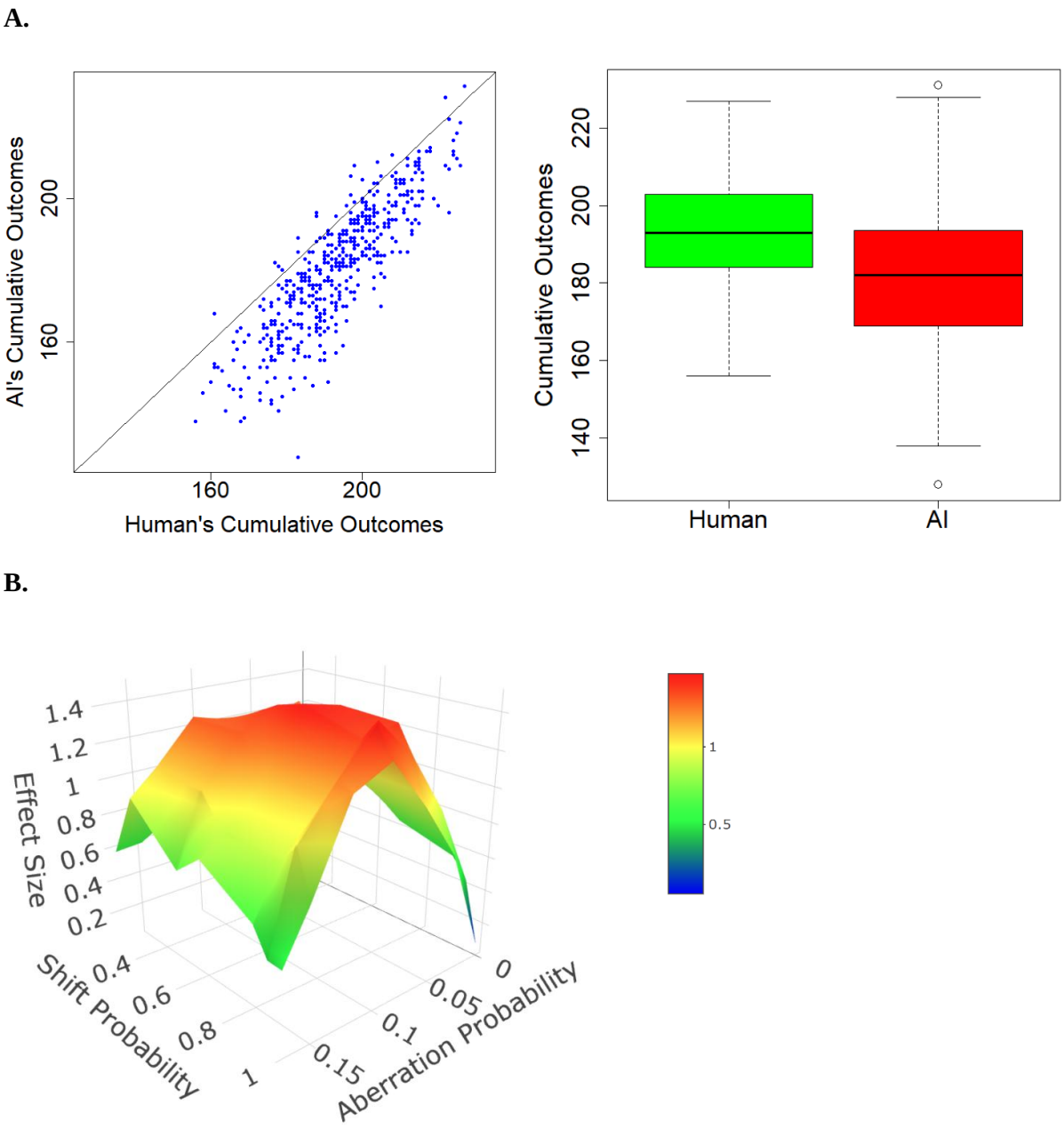


Figure 7. Performance comparison between pattern-aware machines and human forecasters. (**A**) Comparison across 500 Monte Carlo simulations of the current task configuration. **Left**: Scatterplot showing the mean cumulative outcomes earned by humans (x-axis) versus machines (y-axis) in each of the 500 simulations. Points below the 45-degree line indicate superior human performance. **Right**: Boxplot showing the distribution of cumulative outcomes across the 500 task simulations. The humans display higher mean outcomes and lower variance compared to the machines. (**B**) Heatmap showing the effect size (Cohen's d) of the human vs. machine performance differential (paired t-test described in the main text) across various environmental configurations, with shift and aberration probabilities systematically varied across simulation batches.

4 Discussion

We study how and how well people handle a universal cognitive challenge: inferring whether a sudden shock in a system is transient noise or a structural change. Through computational modelling, we show that reducing deep uncertainty through DGP-extraction is an essential prerequisite to navigate this challenge, as applying sophisticated (model-based) thinking to the wrong task structure leads to poor forecasting performance, while applying it to the actual structure provides a significant edge.

Our results demonstrate that participants in the deep uncertainty treatment (T1) successfully overcame both cognitive challenges: they extracted the underlying DGP and applied model-based thinking to it. Furthermore, at least 40% of these participants (likely more[37]) achieved complete DGP extraction, with successful identification of both types of shocks—shifts and aberrations These findings were unexpected, as evidenced by the falsification of our core hypothesis **H1**.[38] They provide the first empirical evidence of the extraction of a complex stochastic DGP, calling for further investigation to replicate and extend these findings. An important direction for future research is understanding the apparent discrepancy between our findings and those of Kendall and Oprea (2024), where participants struggled to extract what appeared to be a simpler DGP. We speculate that this discrepancy reflects three features of our design: (i) substantial performance incentives, (ii) the visual nature of pattern recognition, and (iii) the direct, hands-on nature of the task:

(i) Substantial performance incentives. As explained in the paper, we implemented substantial monetary incentives based on the premise that optimal behavior would be unlikely to emerge otherwise.

(ii) Visual nature of pattern recognition. Visual representations are ubiquitous in domains where shift detection is critical—from geology to ecology to financial economics (e.g., heat maps, spatial correlations, price charts). Yet, many statistical patterns also manifest in numerical formats, and future research should examine whether the pattern extraction ability evidenced in this study generalizes beyond visual displays.

(iii) Direct, hands-on nature of the task. Oprea (2020) highlights that procedural complexity matters alongside computational complexity: mental procedures (e.g., imagining interactions and outcomes) are more complex than physical procedures (executing actions directly). Applied to our setting, this suggests that the pregame phase (purely observational) was more procedurally com-

[37] Those who made occasional mistakes with aberrations (Supplementary Table S4, "Infrequent mistakes") likely extracted the DGP fully but were prone to the complexity-driven mistakes discussed in Section 3.2.4.

[38] The T1 results challenge prevailing assumptions in the field, as reflected in both Ethics Committee concerns about task feasibility (see Footnote 27) and researchers' reactions during seminars (e.g., "It's impossible to detect the pattern from the pregame history"—John List).

plex than the game itself. Similarly, Kendall and Oprea (2024)'s task may have involved greater procedural complexity than our video game-like interface, which provides immediate feedback on actions, creating a hands-on and interactive learning environment—potentially explaining participants' stronger performance in our study. However, evidence that pattern-aware participants had already extracted the DGP during the more procedurally complex pregame phase (Section 3.2.2) challenges this explanation. Further research is needed to explore this question and refine our understanding of what defines cognitive complexity for the human brain.

We also hope that our comparison of machine and human performance serves as a catalyst for further research. Our findings highlight a promising avenue for human-AI collaboration: humans' sophisticated model-based cognition could identify optimal strategies, while machines' consistent execution—less prone to complexity-driven mistakes—could implement them reliably. Future studies could also test AI systems in our experiment: would they detect the pattern in T1? While current evidence suggests that pattern recognition remains challenging for AI, this warrants direct experimental testing. Similarly, would the T2 nudge facilitate AI pattern recognition as effectively as it aided human participants in extracting the task's DGP?

This finding regarding the effectiveness of the T2 nudge in the current experiment demonstrates the potential of information nudges when combined with adequate incentives. While success with information ("*S2*") nudges has been documented, in line with the growing 'active thinker' paradigm, such interventions rarely target DGP learning in complex cognitive tasks—a direction our T2 nudge explores. In this respect, our finding aligns most closely with Graeber (2023), who successfully nudged task participants into attending to the nuisance features of their experimental task's DGP. Further research should examine which kinds of nudges can effectively facilitate DGP learning in complex cognitive tasks and identify the conditions for their success.

A APPENDIX

A.1 Context-specific Reinforcement Learning (RL)

A.1.1 Algorithm variants and fit to human behavior

The 'context-specific RL' model in our main analysis estimates action values using the standard adaptive expectations rule. We consider two types of agents: "Standard RL" learns from experienced outcomes only (Fig. 8A), while "Sophisticated RL" learns from experienced, counterfactual, and observational outcomes (Fig. 8B).

Formally, action value is updated as follows:

$$Q_c(z,t) = \begin{cases} Q_c(z,t-1) + OA(z,t) \cdot LR_c(z,t) \cdot PE(z,t), & \text{if } c = \mathfrak{C}(t), \\ Q_c(z,t-1), & \text{if } c \neq \mathfrak{C}(t). \end{cases} \tag{2}$$

In Equation (2):

- $Q_c(z,t)$ denotes the value of action z in context c, updated after the outcome in trial t. Here, $z \in$ {'No Switch', 'Skip', 'Switch'}, and the set of contexts depends on pattern recognition. In our main analyses, there are 6 contexts for pattern-aware agents (see Section 2.2.1) and 2 for non-recognition agents (see Section 2.2.3).
- $\mathfrak{C}(t)$ indexes the current context that trial t belongs to. Only the Q-values in the current context are updated. The Q-values in other contexts remain unchanged.
- $OA(z,t)$ denotes the outcome appraisal by the agent:
  - $OA(z,t) = 1$ for sophisticated RL agents, indicating that the Q-values of all forecasts are updated, regardless of the actual action taken in trial t.
  - $OA(z,t) = I(z,t)$ for standard RL agents, where $I(z,t)$ is an indicator function that equals 1 if action z is chosen in trial t, and 0 otherwise. In this case, only the Q-value of the selected forecast is updated.
- $LR_c(z,t)$ is the learning rate for context c in trial t.
- $PE(z,t) = \pi(z,t) - Q_{\mathfrak{C}(t)}(z,t-1)$ is the prediction error for action z in trial t, where $\pi(z,t)$ is the outcome from choosing action z in trial t.

In each trial t, the RL agent first identifies the current context of the task, which is $\mathfrak{C}(t)$, and then uses the most recent Q-values in this context to make a forecast. The probability of choosing forecast z in trial t is determined by the following logit rule:

$$P(z_t = z) = \frac{e^{\beta \cdot Q_{\mathfrak{C}(t)}(z,t-1)}}{\sum_z e^{\beta \cdot Q_{\mathfrak{C}(t)}(z,t-1)}}, \tag{3}$$

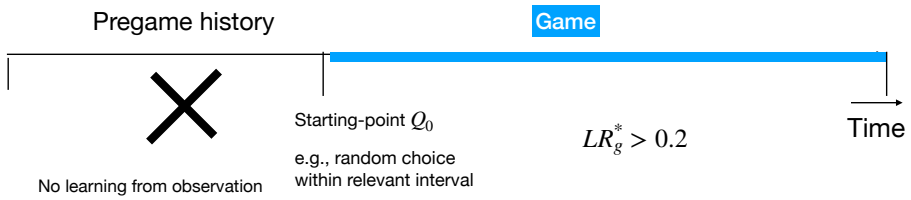
where z_t denotes the forecast chosen at trial t, and β is the logit parameter (larger values imply more deterministic actions—less "trembling"). After generating the forecast z_t and observing the corresponding outcome, the agent updates the Q-value according to Eq.(2) and then proceeds to the next trial $t+1$.

We explore three learning rate scenarios: a single rate across contexts, context-specific rates (6 parameters for the pattern-aware case, 2 for non-recognition), and an endogenous rate that is adjusted on a trial-by-trial basis to reflect recent forecasting performance, where large prediction errors prompt faster learning, akin to the approach described in Pearce and Hall (1980):

$$LR_c\,(z,t) = \frac{2}{1+e^{|PE(z,t)|}} - 1. \tag{4}$$

These variations in outcome appraisal (2) and learning rate (3) yield $2 \times 3 = 6$ model variants.[39] We fitted each to 20 randomly selected 'pattern-aware' and 20 'non-recognition' participants (preliminarily identified based on cumulative outcomes, as shown in Table 1 and Fig. 4 in the main text), following the procedures outlined in Section 2.3.2.

A Standard RL:



B Sophisticated RL:

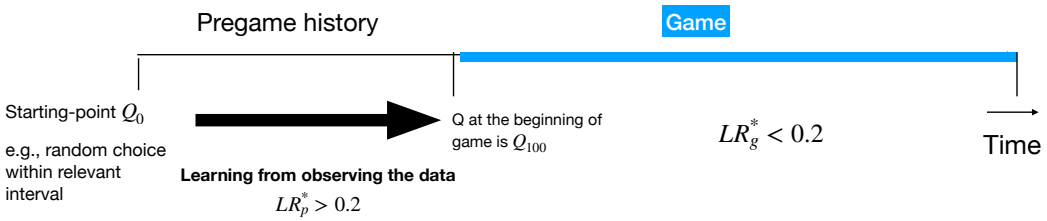


Figure 8. Reinforcement Learning (RL) variants in the task. (A) "Standard RL" learns from experienced outcomes only; as such, it cannot learn from the pregame history. **(B)** "Sophisticated RL" learns equally from experienced, counterfactual, and purely observational outcomes, so the Q value at the beginning of the game is the Q value at the end of the pregame history ("Q_{100}"). LR_p^* and LR_g^* denote the pregame and in-game learning rate values optimized for maximum task earnings as per the procedure described in Appendix A.1.2.

Results show that for the non-recognition group, the standard variant with context-specific learning rates (2 rates, 1 for each context) provides the best fit. For the pattern-aware group, the sophisticated variant with uniform learning rate (1 pregame, 1 in-game) fits best. These winning models represent the 'non-recognition context-specific RL' and 'pattern-aware context-specific RL' types, respectively, in our main analysis.

[39] Simulations confirmed the non-criticality of initial Q-values: using random values within the outcome range (Forecast N: $[-3, 1]$; Forecast S: $[-1, 3]$; Skip: 0) or the average value within the range produced equivalent behaviors.

A.1.2 Optimizing machine learning in the task

We optimized the Reinforcement Learning (RL) model described above for maximum earnings in task simulations, varying outcome appraisal, learning rate, and initial Q value parameters in separate simulation batches.

A counterfactual variant with observational learning in the pregame history (Fig. 9B) significantly outperformed the standard RL algorithm (Fig. 9A). We therefore endowed the machine with these features in the contest documented in Section 3.3.

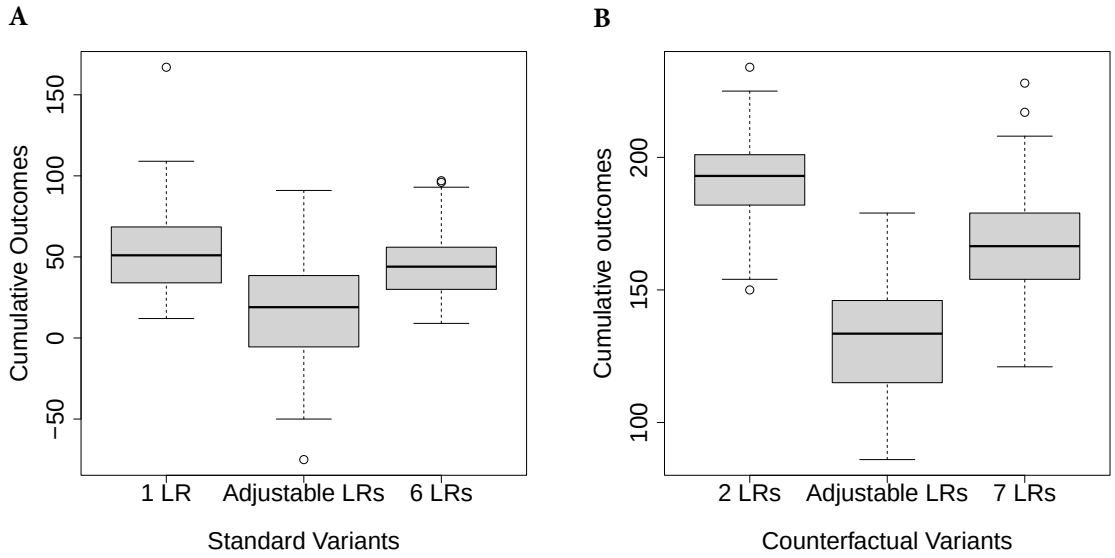


Figure 9. Performance comparison of standard (A) and counterfactual (B) reinforcement learning (RL) variants. Each boxplot shows outcome distributions for 100 test simulations, based on parameter optimization from 100 training simulations. (**A**) Standard RL variants. "1 LR" used a single learning rate, "Adjustable LRs" adapted the rate trial-by-trial based on recent forecasting performance (see Eq. 4), "6 LRs" employed six context-specific rates. (**B**) Counterfactual RL variants. "2 LRs" used two fixed rates (1 pregame, 1 in-game), "Adjustable LRs" same as Panel A, "7 LRs" included one pregame and six context-specific in-game rates.

Fixed learning rate parameters outperformed the Pearce-Hall variant with adjustable rates, as the latter introduced learning instability. Similarly, the version with seven learning rates (one pregame and six in-game, corresponding to each relevant task context) underperformed compared to the simpler two-rate variant (one pregame, one in-game), likely due to the challenges of optimizing a high-dimensional parameter space. Consequently, we equipped the machine with two fixed rates and granted it unlimited pre-task samples to fine-tune these parameter values.

The top-performing variant used a higher pregame learning rate (LR^p in the $[0.2, 0.4]$ range) and a lower in-game learning rate (LR^g in the $[0.1, 0.2]$ range), as shown in Fig. 10. This setup allows for rapid learning from a given starting point (Q_0) and stability during the game. The initial in-game

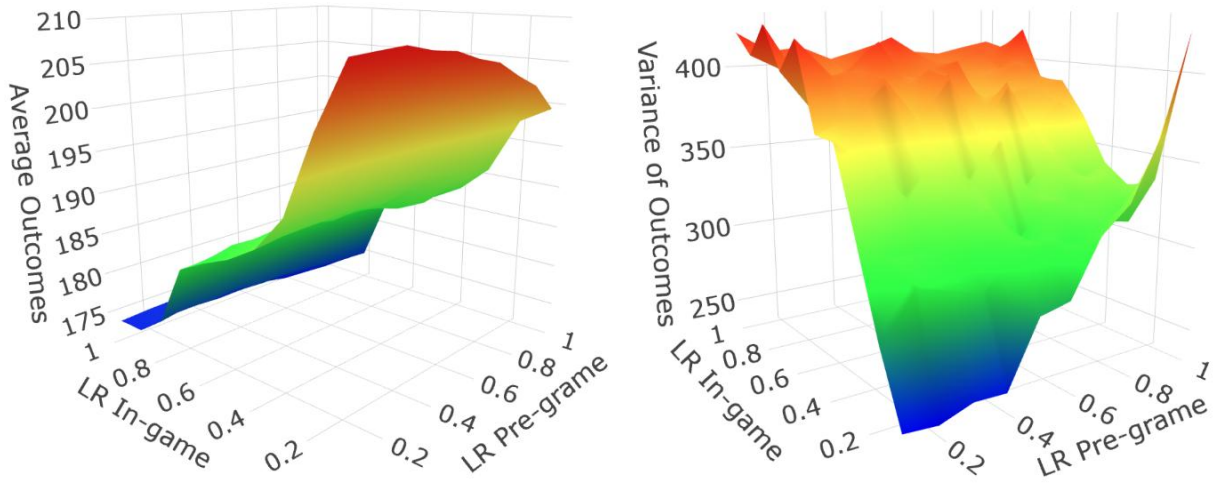


Figure 10. Learning rate optimization for counterfactual Reinforcement Learning (RL). The figure shows the impact of learning rates on the counterfactual RL variant's average cumulative outcomes across 500 simulations. Left: Heatmap of average cumulative outcomes for different learning rates. Right: Heatmap of outcome variance for different learning rates.

Q-value (Q_{100}) often matches the expected value, a result of the extensive pregame history.[40] In such cases, LR_g should ideally be zero for maximum stability. However, unrepresentative pregame endings can bias this value (e.g., the occurrence of a long zone biases the value of the 'No Switch' Forecast upward), requiring a higher LR_g to correct the Q estimate. The optimal LR_g^* balances stability with adjustments for path dependencies. We took this optimized algorithm to represent the machine in the human vs. machine contest documented in Section 3.3.

A.2 EWA

For an EWA agent, the value of each action depends on the past outcomes experienced with this action and is updated as follows:

$$A(z,t) = \frac{\phi N(t-1) A(z,t-1) + [\delta + (1-\delta)\, I(z,t)]\, \pi(z,t)}{N(t)}, \tag{5}$$

where $A(z,t)$ is the value of action $z \in \{$'No Switch', 'Skip', 'Switch'$\}$ updated after experiencing the outcome in trial t, $I(z,t)$ and $\pi(z,t)$ are the functions defined in Eq. (2), and $N(t)$ measures the amount of experience the agent has accumulated after t trials in the game. Starting with an initial value $N(0)$, it is updated as follows:

$$N(t) = (1-\kappa)\, \phi N(t-1) + 1, \qquad t > 1.$$

[40] We found the value of starting point Q_0 has negligible impact on earnings.

$N(0)$, ϕ, κ, and δ are free parameters governing different learning aspects.[41] Our model comparison procedure (Section 2.3.2) involved fitting these parameters along with the initial values $A(\cdot, 0)$, which turn out to have a significant impact on EWA behavior in the present task.

In each trial t, the action probabilities governing the behavior of the EWA agent are determined by $A(\cdot, t-1)$ with the logit rule as in Eq. (3). After generating a forecast and recording trial t outcome(s), the agent updates $A(\cdot, t)$ according to Eq. (5) and then proceeds to the next trial $t+1$.

A.3 Model-based behavior

A model-based agent estimates forecast values through switch probability inference, which we derive using Bayesian principles (without loss of generality—we checked a frequentist approach generates similar behavior in this task). The agent classifies the game trials into C different contexts and estimates switch probability in each context as follows. Consider the probability vector $\boldsymbol{p} = (p_1, p_2, \ldots, p_C) \in [0,1]^C$, where the i-th element represents the switch probability in the i-th context. We discretize the continuous space $[0,1]^C$ into $\mathcal{S}^C$, where $\mathcal{S} = \{0, 0.05, \ldots, 0.95, 1\}$, following d'Acremont and Bossaerts (2016) and without loss of generality.[42] Each element $\boldsymbol{s}$ of the space represents a scenario of the underlying switch probabilities in the C contexts. For example, $\boldsymbol{s} = 0.5^C$ represents the scenario where the switch probabilities are equal to 0.5 in all contexts. We denote by $b(\boldsymbol{s}, t)$ the Bayesian belief assigned to scenario $\boldsymbol{s}$ based on all the information recorded till and including trial t: $b(\boldsymbol{s}, t) = P(\boldsymbol{p} = \boldsymbol{s} \mid \mathcal{F}_t)$, where the filtration $\mathcal{F}_t$ encompasses all the information at the agent's disposal till and including trial t.

To make his decision at trial t, the agent first estimates the switch probability across contexts by taking the weighted average of all scenarios, i.e., $\widehat{\boldsymbol{p}}_{t-1} = \sum_{\boldsymbol{s} \in \mathcal{S}^C} b(\boldsymbol{s}, t-1) \cdot \boldsymbol{s}$, and selecting its $\mathfrak{C}(t)$-th element, which is the switch probability that applies to the current context $\mathfrak{C}(t)$. This switch probability estimate, which we denote by $\hat{p}_{t-1,\mathfrak{C}(t)}$, is used to compute forecast value:

$$\begin{aligned} V(\text{'No Switch'}) &= \hat{p}_{t-1,\mathfrak{C}(t)} \cdot 3^{\alpha} - \lambda\left(1 - \hat{p}_{t-1,\mathfrak{C}(t)}\right), \\ V(\text{'Switch'}) &= 1 - \hat{p}_{t-1,\mathfrak{C}(t)} - \hat{p}_{t-1,\mathfrak{C}(t)} \cdot \lambda \cdot 3^{\alpha}, \end{aligned}$$

where λ and α are the standard loss aversion and risk aversion parameters described by Prospect Theory ($\lambda > 1$ means the agent is loss averse, while $0 < \alpha < 1$ means the agent is risk averse for gains and risk seeking for losses).[43] Based on the above two V-values and $V(\text{'Skip'}) = 0$, the action

[41] A detailed discussion and interpretation of the various EWA parameters can be found in Subsection B.2.1 of Payzan-LeNestour (2018).

[42] This discretization simplifies computations while also reflecting bounded cognitive capacity in humans. Importantly, the degree of coarseness of this discretization does not affect our main results—a result of the extensive pregame history.

[43] We assume the same degree of risk aversion in the gain domain and risk seeking in the loss domain, for simplicity and to avoid scaling issues (Wakker, 2010).

probabilities are determined with the logit rule as in Eq. (3).

At the end of trial t, the agent observes the asset movement realized at trial t. The trial t observation is denoted by a variable $Y_t \in \{-1, 1\}$, with $Y_t = -1$ indicating no switch and $Y_t = 1$ indicating a switch. For a given realization $Y_t = y_t$, the agent updates his beliefs applying Bayes rule:

$$b(\boldsymbol{s}, t) \propto P(Y_t = y_t \mid \boldsymbol{p} = \boldsymbol{s}, \mathcal{F}_{t-1}) \cdot b(\boldsymbol{s}, t-1), \qquad \sum_{\boldsymbol{s} \in \mathcal{S}^C} b(\boldsymbol{s}, t) = 1,$$

where

$$P(Y_t = y_t \mid \boldsymbol{p} = \boldsymbol{s}, \mathcal{F}_{t-1}) = \begin{cases} s_{\mathfrak{C}(t)}, & y_t = 1, \\ 1 - s_{\mathfrak{C}(t)}, & y_t = -1, \end{cases}$$

and $s_{\mathfrak{C}(t)}$ is the $\mathfrak{C}(t)$-th element in $\boldsymbol{s}$, i.e., the switch probability applying to trial t context.

A.3.1 Optimal model

The optimal agent applies the above computations to the six task-relevant contexts described in Section 2.2.1 ($C = 6$). Importantly, its behavior remains consistent across different risk and loss attitudes. This robustness arises because, by design, the estimated shift probability in the dangerous zone is high enough for the agent to consistently forecast 'Switch', regardless of risk preferences (Fig. 11A). While deviations from the risk-neutral benchmark have little impact on overall performance, they increase variance (Fig. 11B). Therefore, we adopted the risk-neutral benchmark in our main analysis for its tractability and lower variance.

A.3.2 Non-recognition model-based type

Non-recognition participants may consider two contexts—a switch has just occurred' and no switch has just occurred' ($C = 2$)—or operate in a context-free manner ($C = 1$). To determine which variant better captures participant behavior, we randomly selected 20 participants classified as 'non-recognition' based on their cumulative outcomes (see Table 1 and Fig. 4A in the main text) and fitted both variants to their behavior, following the procedure outlined in Section 2.3.2. The results indicate that the context-free variant better fits the data. Accordingly, we used this variant to represent the non-recognition model-based type in our main analysis.

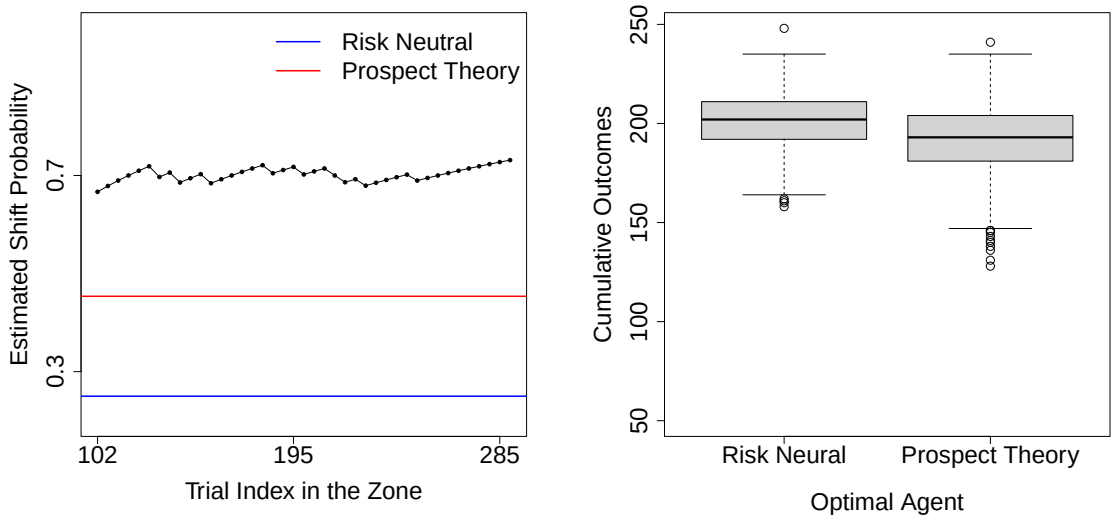


Figure 11. Robustness of optimal behavior to varying risk preferences. (**A**) Estimated shift probability in the zone during a simulated task run by an optimal forecaster. Horizontal lines indicate the Switch' forecast thresholds: black for a risk-neutral agent, and red for the standard Prospect Theory agent ($\lambda = 2, \alpha = 0.8$). Both agents make deterministic choices (high logit parameter) and forecast Switch' when the estimated probability surpasses the corresponding threshold.(**B**) Cumulative outcomes for risk-neutral and standard Prospect Theory optimal forecasters across the 72 lab data series. In these simulations, the logit parameter was set to the mean fitted value across all participants in our lab experiment.

INTERNET APPENDIX (NOT FOR PUBLICATION)

Online Appendix A Performance edge of optimal forecasters

We simulated the task 500 times for each of the five models described in Appendix A.

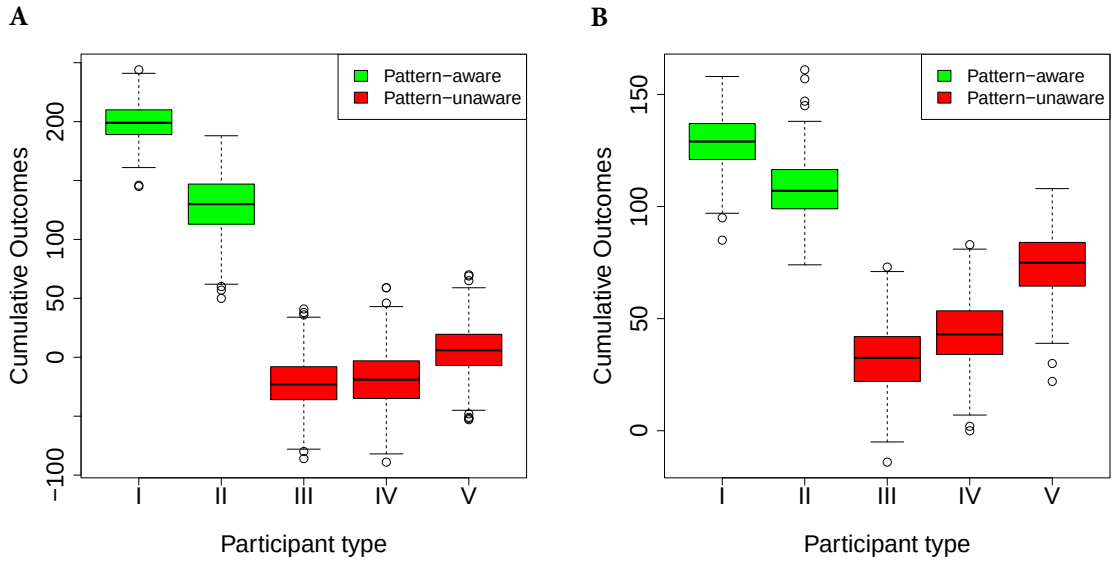


Figure S12. Performance comparison of the five participant types. Labels I to V correspond to the optimal, pattern-aware RL, non-recognition RL, EWA, and non-recognition model-based types, as defined in Section 2.2 of the main text. **(A)** Cumulative outcome distributions from 500 simulations of the task for each type. **(B)** Cumulative outcome distributions from 500 simulations of a task variant with symmetric payoffs (1 for correct forecasts, −1 for incorrect ones), for each type.

Fig. S12A shows that the optimal model achieves an average cumulative outcome of \$208 (SEM: 1.7) across 500 Monte Carlo simulations of the task, significantly outperforming the second-best performer ('pattern-aware context-specific RL' type: \$135; SEM: 3). A paired t-test confirms this outperformance, rejecting the null hypothesis that the optimal approach earns less ($t = 34.405$, $p < 0.001$, one-tailed; Cohen's D = 4.1). The optimal approach also demonstrates superior stability, with lower variance in outcomes. Fig. S12B shows that this performance advantage diminishes markedly in simulations of a task variant with symmetric payoffs.[44]

[44]These results are not sensitive to chosen parameter values. For Fig S12: Adaptive models used mid-range initial values (−1 for 'N', 1 for 'S'). EWA model had $N(0) = 10$, other learning parameters were set to 0.5. Pattern-unaware RL and Pattern-aware RL used 1 for learning rates, consistent with prior economics and neuroscience studies (e.g., Evans and Honkapohja (2001), Behrens et al. (2007), and MMcGuire et al. (2014)) and without loss of generality: RL's underperformance compared to the optimal model is observed across all learning rates, see Section 3.3. For all models, the logit parameter was set to 4.4, which is the average fitted value across participants.

Online Appendix B Recovery, power, and robustness of model comparison

We conducted extensive analyses to assess the recovery properties of our model comparison procedure (described in Section 2.3.2). First, we checked our ability to discriminate between the optimal and best-fit non-recognition models in Tests 1, 2, and 4 (as defined in Section 2.3.2). Simulating 72 optimal agents, each assigned a lab experiment data series (Section 3.1), showed the optimal model fits best for all, indicating a 100% recovery rate. We also found 100% recovery when simulating 72 non-recognition agents (Figs. 2-3 in our preregistration file [here]).

We further assessed the power of these tests by running a simulated experiment with 70% optimal forecasters and 30% non-recognition RL agents ($N = 70$). The out-of-sample accuracy criterion allowed us to reject the null hypothesis that the RL model fits better ($t = 1.69$, $p = 0.05$; Cohen's D= 0.2), but the log-likelihood criterion did not ($t = 0.3$, $p = 0.38$; Cohen's D= 0.04). We therefore used the accuracy criterion for Tests 1, 2, and 4.

Next, we evaluated the recovery property of our procedure to discriminate between optimal and pattern-aware RL models in our follow-up paired t-test (Section 2.3.2). Comparing the two models' goodness-of-fit in simulations of 72 optimal forecasters (Fig. S13A) and 72 pattern-aware RL agents (Fig. S13B) yielded recovery rates exceeding 95%. We further evaluated the power of this test by running a simulated experiment with 70% optimal forecasters and 30% pattern-aware RL agents ($N = 70$). We could reject the null hypothesis using either the accuracy or log-likelihood metric. (N.B: We constrained the RL model's learning rate above 0.4 both for plausibility and recovery purposes.)

Robustness checks. As noted, both accuracy and log-likelihood (LL) metrics can be used for the follow-up paired t-test described in the main text. To test the robustness of our conclusions, we applied the test separately with each metric. Similarly, we tested the robustness of Tests 3 and 5 by re-running them using both accuracy and LL criteria to classify participants. When the criteria disagreed or the goodness-of-fit difference between the best and second-best models was under 5%, we referred to end-of-game debriefings: participants describing the pattern were classified as pattern aware, those who dismissed the indicator chart as useless were classified as pattern unaware, and the rest were excluded (1% for T1, 3% for T2, 5% for T3). The results reported in the main text remain qualitatively unchanged and are further strengthened under this procedure.

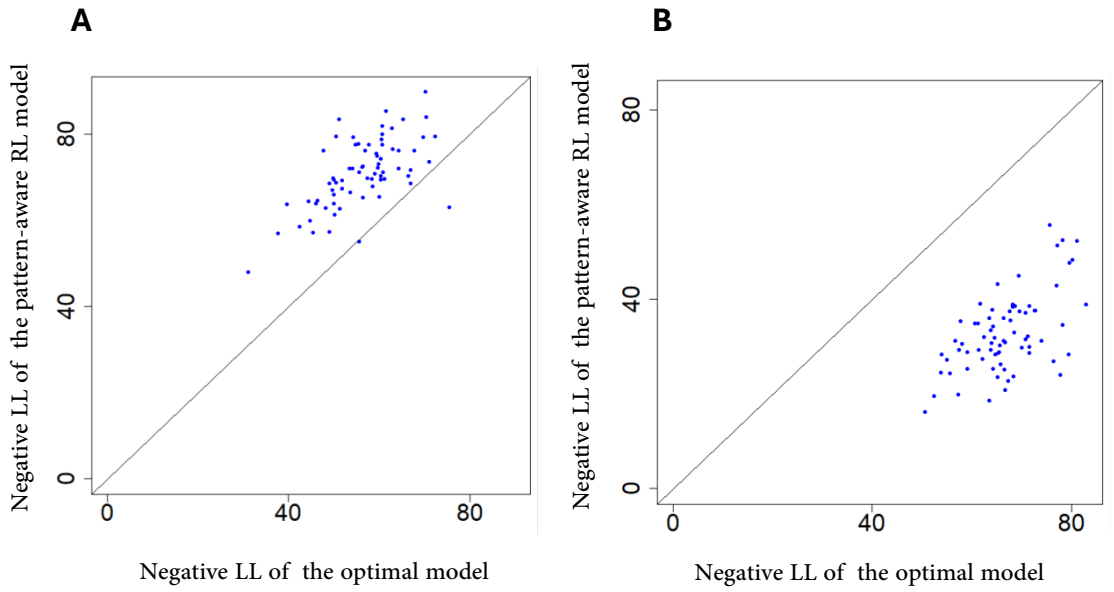


Figure S13. Discriminatory power for optimal vs. pattern-aware RL models. Each panel shows the out-of-sample negative log-likelihood (LL) of the optimal model (x-axis) versus the pattern-aware RL model (y-axis) in a simulated experiment with 72 optimal forecasters (**A**) and 72 pattern-aware RL agents (**B**). Each data point represents one simulated agent. The optimal model fits better when the data point is above the 45-degree line.

Online Appendix C Supplementary figure

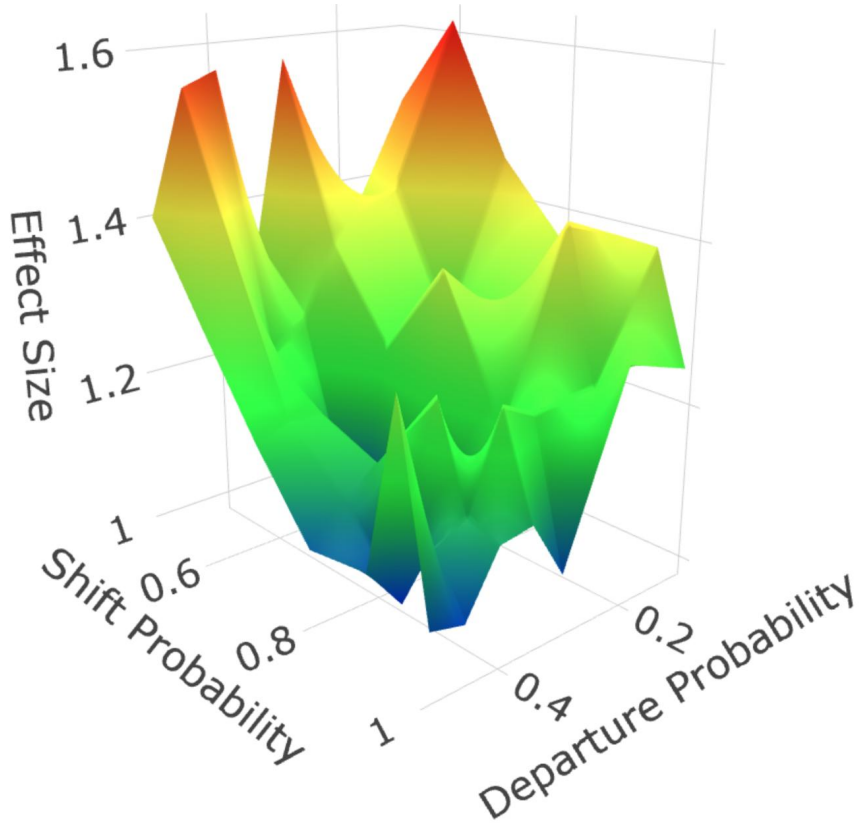


Figure S14. Heatmap showing the effect size (Cohen's d) of the human vs. AI performance differential (paired t-test described in Section 3.3 of the main text) across various environmental configurations, varying shift and departure probabilities in the experimental task.

Online Appendix D Supplementary statistics

Table S2 presents participant demographics for gender, age, and academic discipline, across the treatment groups. The table shows balanced proportions for each characteristic, ensuring comparability between groups.

Table S3 presents participant performance across treatments. The average click frequency is significantly lower in T1 compared to other treatments, driven by a significant fraction (20%) of participants who never clicked. Fig. S15 shows that T1 participants' click frequency distribution is bimodal and right-skewed, with the majority clicking close to 100% of the time.

Fig. S16 displays the performance of the 'Higher Performers' and 'Lower Performers' groups as defined in the main text. The 'Higher Performers' group clusters in the right corner of the graph, demonstrating strong performance in both normal times and switch events, while the 'Lower Performers' group shows poor performance in at least one dimension.

Table S4 shows the distribution of pattern-aware participants' responses to aberration events, categorized by their frequency of mistakes—defined as forecasting 'No Switch' during these self-correcting shocks.

Table S2. Summary of Participant Demographics by Treatment Group

	Gender			Age			Discipline		
	F	M	Other	<20	20-25	>25	Business	STEM	Other
T1	25	46	0	8	57	6	64	5	2
T2	23	39	0	8	49	5	57	3	2
T3	22	39	0	13	35	13	52	5	1
N	70	124	0	29	141	24	173	13	5

Table S3. Summary of Participant Performance by Treatment Group. Numbers represent means, with standard deviation in parenthesis for cumulative outcomes.

	T1 (N=71)	T2 (N=62)	T3 (N=60)
Performance variables:			
Cumulative outcomes ($)	86 (64)	130 (56)	154 (59)
Click (%)	68	93	97
Skip (%)	14	9	5
Missed (%)	<0.1	<0.1	<0.1
Response time (s)	6.5	2.4	2.5

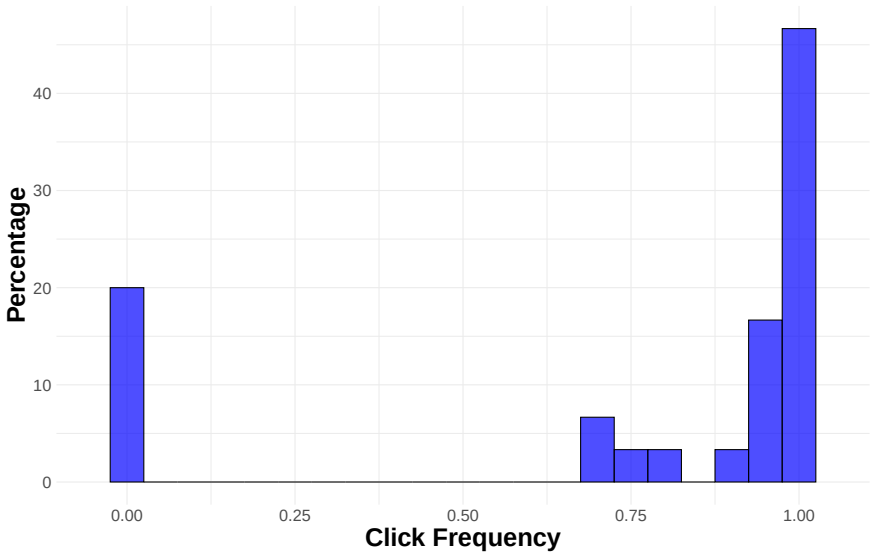


Figure S15. Distribution of T1 participants' average click frequency throughout the task.

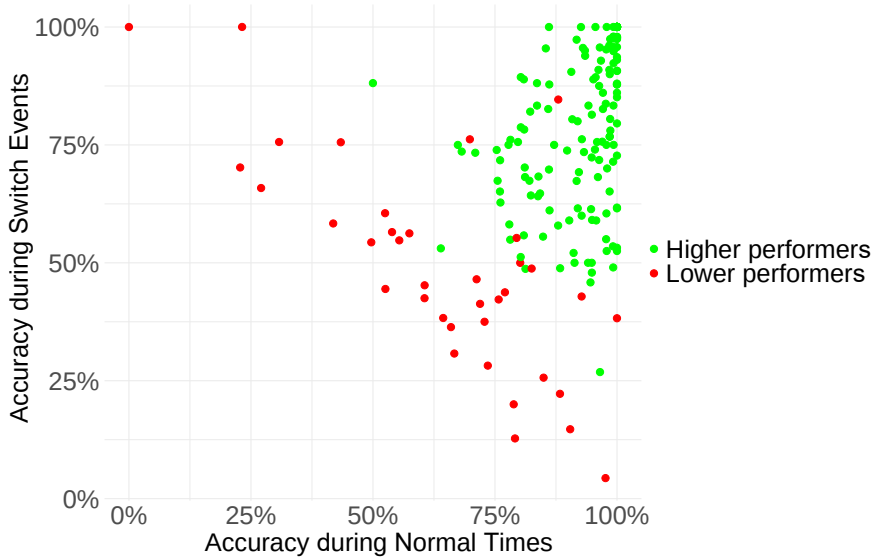


Figure S16. Forecast accuracy of the lab participants. Forecast accuracy measured during normal times (x-axis) and switch events (y-axis). Data pools participants from all three treatments, with green dots representing Higher Performers and red dots Lower Performers as defined in the main text.

Table S4. Distribution of participant responses to aberration events. Percentages shown are calculated among pattern-aware participants for each treatment condition. "Ignored aberrations": participants who selected 'No Switch' for $\geq$ 90% of aberration events (systematic error); "Frequent mistakes": participants who selected 'No Switch' for 50-90% of aberration events. "Infrequent mistakes": participants who selected 'No Switch' for 20-50% of aberration events; "Identified aberrations": participants who only made occasional errors ($\leq$ 20% 'No Switch' responses).

	Percentage among pattern-aware participants (%)			
	Ignored aberrations	**Frequent mistakes**	**Infrequent mistakes**	**Identified aberrations**
Treatment 1	11	28	22	39
Treatment 2	10	21	29	40
Treatment 3	2	11	20	67

References

Acemoglu, D. and Restrepo, P., 2018. The race between man and machine: Implications of technology for growth, factor shares, and employment. American Economic Review, 108(6), pp.1488-1542.

Acharya, V.V., Johnson, T., Sundaresan, S. and Zheng, S., 2023. Disasters with unobservable duration and frequency: Intensified responses and diminished preparedness (No. w31067). National Bureau of Economic Research.

Afrouzi, H., Kwon, S.Y., Landier, A., Ma, Y. and Thesmar, D., 2023. Overreaction in expectations: Evidence and theory. The Quarterly Journal of Economics, 138(3), pp.1713-1764.

Aher, G.V., Arriaga, R.I. and Kalai, A.T., 2023, July. Using large language models to simulate multiple humans and replicate human subject studies. In International Conference on Machine Learning (pp. 337-371). PMLR.

Alós-Ferrer, C. and Garagnani, M., 2023. Part-time Bayesians: incentives and behavioral heterogeneity in belief updating. Management Science, 69(9), pp.5523-5542.

Aragones, E., Gilboa, I., Postlewaite, A. and Schmeidler, D., 2005. Fact-free learning. American Economic Review, 95(5), pp.1355-1368.

Asparouhova, E., Hertzel, M. and Lemmon, M., 2009. Inference from streaks in random outcomes: Experimental evidence on beliefs in regime shifting and the law of small numbers. Management Science, 55(11), pp.1766-1782.

Ba, C., 2021. Robust Misspecified Models and Paradigm Shifts. arXiv preprint arXiv:2106.12727.

Barberis, N.C. and Jin, L.J., 2023. Model-free and model-based learning as joint drivers of investor behavior (No. w31081). National Bureau of Economic Research.

Barberis, N., Shleifer, A. and Vishny, R., 1998. A model of investor sentiment. Journal of financial economics, 49(3), pp.307-343.

Behrens, T.E., Woolrich, M.W., Walton, M.E. and Rushworth, M.F., 2007. Learning the value of information in an uncertain world. Nature Neuroscience, 10(9), pp.1214-1221.

Benjamin, D.J., 2019. Errors in probabilistic reasoning and judgment biases. Handbook of Behavioral Economics: Applications and Foundations 1, 2, pp.69-186.

Bordalo, P., Conlon, J.J., Gennaioli, N., Kwon, S.Y. and Shleifer, A., 2023. How people use statistics (No. w31631). National Bureau of Economic Research.

Bossaerts, P., 2018. Formalizing the function of anterior insula in rapid adaptation. Frontiers in Integrative Neuroscience, 12, p.61.

Brav, A. and Heaton, J.B., 2002. Competing theories of financial anomalies. The Review of Financial Studies, 15(2), pp.575-606.

Camerer, C. and Hua Ho, T., 1999. Experience-weighted attraction learning in normal form games. Econometrica, 67(4), pp.827-874.

Cao, S., Jiang, W., Wang, J. and Yang, B., 2024. From man vs. machine to man+ machine: The art and AI of stock analyses. Journal of Financial Economics, 160, p.103910.

Charles, C., Frydman, C. and Kilic, M., 2022. Insensitive investors. The Journal of Finance, 79(4), pp.2473-2503.

Charness, G. and Levin, D., 2005. When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. American Economic Review, 95(4), pp.1300-1309.

Charness, G., Karni, E. and Levin, D., 2010. On the conjunction fallacy in probability judgment: New experimental evidence regarding Linda. Games and Economic Behavior, 68(2), pp.551-556.

Chater, N. and Loewenstein, G., 2023. The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. Behavioral and Brain Sciences, 46, p.e147.

d'Acremont, M. and Bossaerts, P., 2016. Neural mechanisms behind identification of leptokurtic noise and adaptive behavioral response. Cerebral Cortex, 26(4), pp.1818-1830.

Dal Bó, P. and Fréchette, G.R., 2011. The evolution of cooperation in infinitely repeated games: Experimental evidence. American Economic Review, 101(1), pp.411-429.

Daw, N.D., 2014. Advanced reinforcement learning. Neuroeconomics, pp.299-320.

Daw, N.D., Niv, Y. and Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neuroscience, 8(12), pp.1704-1711.

Doll, B.B., Simon, D.A. and Daw, N.D., 2012. The ubiquity of model-based reinforcement learning. Current Opinion in Neurobiology, 22(6), pp.1075-1081.

Dunne, S., D'Souza, A. and O'Doherty, J.P., 2016. The involvement of model-based but not model-free learning signals during observational reward learning in the absence of choice. Journal of Neurophysiology, 115(6), pp.3195-3203.

Embrechts, P., Klüppelberg, C. and Mikosch, T., 2013. Modelling Extremal Events: for Insurance and Finance (Vol. 33). Springer Science & Business Media.

Enke, B. and Graeber, T., 2023. Cognitive uncertainty. The Quarterly Journal of Economics, 138(4), pp.2021-2067.

Evans, G.W. and Honkapohja, S., 2001. Expectations, Economics of. In: N.J. Smelser and P.B. Baltes, eds. International Encyclopedia of the Social & Behavioral Sciences. Amsterdam: Elsevier, pp. 5060–5067. doi:10.1016/B0-08-043076-7/02245-2.

Evans, J.S.B. and Stanovich, K.E., 2013. Dual-process theories of higher cognition: Advancing the debate. Perspectives on Psychological Science, 8(3), pp.223-241.

Fan, T., Liang, Y. and Peng, C., 2021. The inference-forecast gap in belief updating. Working paper.

Fenton-O'Creevy, M., Soane, E., Nicholson, N. and Willman, P., 2011. Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. Journal of Organizational Behavior, 32(8), pp.1044-1061.

Findling, C., Chopin, N. and Koechlin, E., 2021. Imprecise neural computations as a source of adaptive behaviour in volatile environments. Nature Human Behaviour, 5(1), pp.99-112.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Oxford: Clarendon Press.

Frydman, C. and Nave, G., 2017. Extrapolative beliefs in perceptual and economic decisions: Evidence of a common mechanism. Management Science, 63(7), pp.2340-2352.

Fudenberg, D., Gao, W. and Liang, A., 2023. How flexible is that functional form? Quantifying the restrictiveness of theories. Review of Economics and Statistics, pp.1-50.

Gabaix, X., 2014. A sparsity-based model of bounded rationality. The Quarterly Journal of Economics, 129(4), pp.1661-1710.

Gabaix, X., 2019. Behavioral inattention. In Handbook of behavioral economics: Applications and foundations 1 (Vol. 2, pp. 261-343). North-Holland.

Gagnon-Bartsch, T., Rabin, M. and Schwartzstein, J., 2018. Channeled attention and stable errors. Boston: Harvard Business School.

Gawdat, M., 2021. Scary Smart: The Future of Artificial Intelligence and How You Can Save Our World. Pan Macmillan.

Gigerenzer, G. and Todd, P.M., 1999. Simple Heuristics That Make Us Smart. Oxford University Press, USA.

Giglio, S., Maggiori, M., Stroebel, J. and Utkus, S., 2021. Five facts about beliefs and portfolios. American Economic Review, 111(5), pp.1481-1522.

Graeber, T., 2023. Inattentive inference. Journal of the European Economic Association, 21(2), pp.560-592.

Hallsworth, M., 2022. Making sense of the "Do nudges work?" debate. Behavioral Scientist, 2.

Handel, B. and Schwartzstein, J., 2018. Frictions or mental gaps: What's behind the information we (don't) use and when do we care?. Journal of Economic Perspectives, 32(1), pp.155-178.

Hasanhodzic, J., Lo, A.W. and Viola, E., 2019. What do humans perceive in asset returns?. Journal of Portfolio Manage, 45, pp.49-60.

Hayes, W.M., Yax, N. and Palminteri, S., 2024. Large language models are biased reinforcement learners. arXiv preprint arXiv:2405.11422.

Hayes, W.M., Yax, N. and Palminteri, S., 2024. Relative value biases in large language models. arXiv preprint arXiv:2401.14530.

He, S. and Kučinskas, S., 2024. Expectation formation with correlated variables. The Economic Journal, 134(660), pp.1517-1544.

Hertwig, R. and Grüne-Yanoff, T., 2017. Nudging and boosting: Steering or empowering good decisions. Perspectives on Psychological Science, 12(6), pp.973-986.

Hertwig, R. and Ortmann, A., 2001. Experimental practices in economics: A methodological challenge for psychologists?. Behavioral and brain sciences, 24(3), pp.383-403.

Ho, T.H., Camerer, C.F. and Chong, J.K., 2007. Self-tuning experience weighted attraction learning in games. Journal of Economic Theory, 133(1), pp.177-198.

Ho, T.H., Wang, X. and Camerer, C.F., 2008. Individual differences in EWA learning with partial payoff information. The Economic Journal, 118(525), pp.37-59.

Hong, H., Stein, J.C. and Yu, J., 2007. Simple forecasts and paradigm shifts. The Journal of Finance, 62(3), pp.1207-1242.

Horton, J.J., 2023. Large language models as simulated economic agents: What can we learn from homo silicus? (No. w31122). National Bureau of Economic Research.

Jones, C.I., 2023. The ai dilemma: Growth versus existential risk (No. w31837). National Bureau of Economic Research.

Kagel, J.H. and Roth, A.E. eds., 2020. The handbook of experimental economics, volume 2. Princeton university press.

Kahneman, D. and Egan, P., 2011. Thinking, fast and slow: Farrar. Straus and Giroux, 1, p.1025.

Kendall, C. and Oprea, R., 2024. On the complexity of forming mental models. Quantitative Economics, 15(1), pp.175-211.

Knight, F.H., 1921. Risk, uncertainty and profit. Vol. 31. Houghton Mifflin.

Kremer, M., Moritz, B. and Siemsen, E., 2011. Demand forecasting behavior: System neglect and change detection. Management Science, 57(10), pp.1827-1843.

Kuhnen, C.M., 2015. Asymmetric learning from financial information. The Journal of Finance, 70(5), pp.2029-2062.

Larson, E.J., 2021. The Myth of Artificial Intelligence: Why Computers Can't Think The Way We Do. Harvard University Press.

Lee, S.W., Shimojo, S. and O'doherty, J.P., 2014. Neural computations underlying arbitration between model-based and model-free learning. Neuron, 81(3), pp.687-699.

Lo, A.W. and Repin, D.V., 2002. The psychophysiology of real-time financial risk processing. Journal of cognitive neuroscience, 14(3), pp.323-339.

Lo, A.W., Repin, D.V. and Steenbarger, B.N., 2005. Fear and greed in financial markets: A clinical study of day-traders. American Economic Review, 95(2), pp.352-359.

Malmendier, U. and Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk taking?. The Quarterly Journal of Economics, 126(1), pp.373-416.

Mandelbrot, B., 1963. The variation of certain speculative prices. The Journal of Business, 36(4), pp.394-419.

Mertens, S., Herberz, M., Hahnel, U.J. and Brosch, T., 2022. The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. Proceedings of the National Academy of Sciences, 119(1), p.e2107346118.

McGuire, J.T., Nassar, M.R., Gold, J.I. and Kable, J.W., 2014. Functionally dissociable influences on learning rate in a dynamic environment. Neuron, 84(4), pp.870-881.

Meyniel, F., Schlunegger, D. and Dehaene, S., 2015. The sense of confidence during probabilistic learning: A normative account. PLoS Computational Biology, 11(6), p.e1004305.

Oprea, R., 2020. What makes a rule complex?. American economic review, 110(12), pp.3913-3951.

Oprea, R., 2024. Decisions under risk are decisions under complexity. American Economic Review, 114(12), pp.3789-3811.

Packard, M.D., Bylund, P.L. and Clark, B.B., 2021. Keynes and Knight on uncertainty: Peas in a pod or chalk and cheese?. Cambridge Journal of Economics, 45(5), pp.1099-1125.

Palminteri, S. and Lebreton, M., 2022. The computational roots of positivity and confirmation biases in reinforcement learning. Trends in Cognitive Sciences, 26(7), pp.607-621.

Palminteri, S., Wyart, V. and Koechlin, E., 2017. The importance of falsification in computational cognitive modeling. Trends in cognitive sciences, 21(6), pp.425-433.

Payzan-LeNestour, E., 2018. Can people learn about 'black swans'? Experimental evidence. The Review of Financial Studies, 31(12), pp.4815-4862.

Payzan-LeNestour, E. and Bossaerts, P., 2011. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. PLoS Computational Biology, 7(1), p.e1001048.

Payzan-LeNestour, E. and Bossaerts, P., 2015. Learning about unstable, publicly unobservable payoffs. The Review of Financial Studies, 28(7), pp.1874-1913.

Payzan-LeNestour, E. and Doran, J., 2024. Craving money? Evidence from the laboratory and the field. Science Advances, 10(2), p.eadi5034.

Pearce, J.M. and Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological review, 87(6), p.532.

Pouget, S., 2007. Adaptive traders and the design of financial markets. The Journal of Finance, 62(6), pp.2835-2863.

Samuelson, W. and Zeckhauser, R., 1988. Status quo bias in decision making. Journal of Risk and Uncertainty, 1, pp.7-59.

Shmaya, E. and Yariv, L., 2016. Experiments on decisions under uncertainty: A theoretical framework. American Economic Review, 106(7), pp.1775-1801.

Simmons, J., Nelson, L. and Simonsohn, U., 2022. Meaningless Means# 1: The Average Effect of Nudging Is d=. 43. Available at: https://datacolada.org/105.

Sunstein, C.R., 2016. The Ethics of Influence: Government in the Age of Behavioral Science. Cambridge University Press.

Szollosi, A. and Newell, B.R., 2020. People as intuitive scientists: Reconsidering statistical explanations of decision making. Trends in Cognitive Sciences, 24(12), pp.1008-1018.

Szollosi, A., Wang-Ly, N. and Newell, B., 2023. Nudges for People who Think. Working Paper. PsyArXiv, June 8. Available at: https://doi.org/10.31234/osf.io/ravsp.

Thorndike, E.L., 1927. The Law of Effect. The American Journal of Psychology, 39, 212-222.

Vieider, F.M., 2024. Decisions under uncertainty as bayesian inference on choice options. Management Science.

Wakker, P.P., 2010. Prospect theory: For risk and ambiguity. Cambridge university press.

Webb, M., 2019. The impact of artificial intelligence on the labor market. Available at SSRN 3482150.

Wilcox, N.T., 1993. Lottery choice: Incentives, complexity and decision time. The Economic Journal, 103(421), pp.1397-1417.

World Bank, 2014. World development report 2015: Mind, society, and behavior. The World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/20597.

Yax, N., Anlló, H. and Palminteri, S., 2024. Studying and improving reasoning in humans and machines. Communications Psychology, 2(1), p.51.

Zizzo, D.J., 2010. Experimenter demand effects in economic experiments. Experimental Economics, 13, pp.75-98.