

# Probabilistic Forecasting Guides Dynamic Decisions

Shuze Liu<sup>1</sup>, Yang Xiang<sup>2</sup>, and Samuel J. Gershman<sup>2, 3, 4</sup>

<sup>1</sup> Program in Neuroscience, Harvard University

<sup>2</sup> Department of Psychology, Harvard University

<sup>3</sup> Center for Brain Science, Harvard University

<sup>4</sup> Center for Brains, Minds, and Machines, Massachusetts Institute of Technology

Real-world assets—such as projects, jobs, skills, or relationships—change endogenously with continued investment, often improving at variable rates. These structural dynamics pose a challenge: How do we decide what asset to invest in and when to divest from an underperforming asset? In this article, we propose a normative model of asset selection and switching. We hypothesize that people use Bayesian function learning to infer invariant, structural parameter underlying changes in asset performance, and choose assets based on forecasted final performances at a task-specific time horizon. We test the model in four experiments ( $N = 460$ ) that manipulate the setup (selection vs. switching), time horizons, and statistical properties of asset performance trajectories. The Bayesian forecasting model consistently outperforms myopic alternatives inspired by foraging and study time allocation, which instead assume direct continuation of recent observations and rely solely on them. Our results highlight a distinctive form of future-looking cognition: Humans anticipate how their own actions will shape future asset performance, and draw on representations of underlying structural dynamics to inform their current choices. The Bayesian probabilistic forecasting model therefore offers a principled computational account of how humans make decisions with lasting consequences.

**Keywords:** decision making, function learning, asset selection, asset switching, learning trajectories

Imagine yourself working on a final course project: One topic is incremental but easy to make progress on, while another is ambitious and slower to yield results. What would you choose if the project were due tomorrow versus in a month? And what if you had already made progress on one topic—would you persist or switch when you feel stuck? Similar trade-offs are pervasive in everyday life, from deciding which job to pursue or research direction to follow to selecting which trainee to hire or whether to continue a relationship. These decisions

carry long-term consequences: Projects, careers, and people typically improve with sustained investment (Topel & Ward, 1992), and such investments are often mutually exclusive—committing to one option precludes simultaneous investment in others. Furthermore, switching to a new alternative is often costly, as doing so forfeits the progress already made. Given the many domains in which similar dynamics arise, we refer to these entities collectively as “assets,” emphasizing that each has a measurable level of performance that evolves with continued investment over time.

The above everyday examples illustrate a broad class of decisions that we refer to as *asset selection and switching* problems. These problems share several features highlighted above: They unfold over finite time horizons, assets exhibit systematic performance dynamics with continued investment, progress is typically nontransferable across assets, and mutual exclusivity prevents simultaneous investment in multiple alternatives. Together, these constraints make asset-related decisions especially challenging: Optimal decision makers must not only evaluate current assets’ performances but also anticipate how fast each asset will develop, how much time is left, and how much progress has already been made. How humans navigate these challenges thus becomes an important yet inadequately explored problem.

Although a variety of fields—such as reinforcement learning (RL), patch foraging, and study time allocation—have examined related dynamic decision problems (Charnov, 1976; Metcalfe & Kornell, 2005; Sutton & Barto, 2018), their standard assumptions place important limits on their applicability to real-world asset selection and switching. Namely, they assume static/stationary components in the environment that enable predicting the future directly from the past (Li et al., 2025). While some form of stationarity is necessary for learning to occur, these frameworks typically assume low-level forms of stationarity that are quite restrictive. In particular, raw observations are treated as Markovian and sufficient for prediction: Future

Peter D. Kvam served as action editor.

Shuze Liu  <https://orcid.org/0000-0002-3946-5003>

The online task, data, and analysis code for all experiments are available on GitHub at [https://github.com/LSZ2001/forecasting\\_guides\\_dynamic\\_decisions](https://github.com/LSZ2001/forecasting_guides_dynamic_decisions). The tasks and analyses have been preregistered on AsPredicted at <https://aspredicted.org/3jv3-kgx.pdf>, <https://aspredicted.org/f33z-s27j.pdf>, <https://aspredicted.org/sh73-d3z4.pdf>, and <https://aspredicted.org/c8nz6a.pdf>. The authors did not publish their results at an earlier point.

The authors have no conflicts of interest to disclose. This work was supported by the Kempner Institute for the Study of Natural and Artificial Intelligence. The authors are grateful to Samuel Hall-McMaster, Arthur Prat-Carrabin, and Tejas Ramdas for helpful discussions.

Shuze Liu played a lead role in data curation, formal analysis, investigation, project administration, validation, visualization, and writing—original draft and an equal role in conceptualization, methodology, and software. Yang Xiang played a supporting role in visualization and an equal role in conceptualization, methodology, software, and writing—review and editing. Samuel J. Gershman played a lead role in funding acquisition and supervision and an equal role in conceptualization, methodology, and writing—review and editing.

Correspondence concerning this article should be addressed to Shuze Liu, Program in Neuroscience, Harvard University, 52 Oxford Street, Cambridge, MA 02138, United States. Email: [shuzeliu@fas.harvard.edu](mailto:shuzeliu@fas.harvard.edu)

observations (e.g., an option's future value) are assumed to depend only on the most recent observation and action (Sutton & Barto, 2018). As a result, there is no need to track elapsed time or incorporate earlier observations and actions into the predictive process (Metcalfe & Kornell, 2005; Sutton & Barto, 2018). Such assumptions allow for elegant and tractable models (e.g., myopic foraging rules are justified when rewards monotonically decline with continued exploitation; Charnov, 1976; Constantino & Daw, 2015), but fail to capture the richer temporal dynamics that characterize real-world asset selection and switching.

To illustrate, consider a variant of the classical foraging problem. In standard foraging, an agent repeatedly reaps rewards by exploiting particular locations ("patches") for food. Continued exploitation gradually depletes a patch, causing its reward to decline over time (Charnov, 1976). In contrast, suppose that continued investment instead increases a patch's value—for example, as in asset investment or skill development, where sustained effort can produce accelerating growth. In such settings, the value of staying or leaving depends on how current actions reshape future performance, rather than on past rewards alone. The original marginal value theorem (MVT; Charnov, 1976), myopically tracking the most recent rewards but not the growth trajectory, is therefore ill-suited to capture these endogenous, action-dependent dynamics (Constantino & Daw, 2015).

As alluded to by the examples above, the world is rarely Markovian at the level of low-level observations. Instead, it is Markovian only with respect to latent structural dynamics—the growth, decay, or learning processes that govern how assets evolve with continued investment. One possible remedy is to modify existing models so that they explicitly track the full history of past actions and observations, allowing the system to retain the Markovian property even when relying only on observable information. However, this strategy quickly becomes intractable for resource-limited agents to learn in finite data samples or to execute in finite time (Sutton & Barto, 2018). A more promising approach is to formulate models built around simple but expressive dynamical structures that capture the relevant forms of endogeneity. Representing these latent structures provides a tractable basis for inference and prediction and thus affords human usage when facing decisions with long-term consequences (S. J. Gershman & Niv, 2015; S. Gershman et al., 2010, 2017).

In summary, the above reasoning highlights a fundamental contrast between *past-looking* and *future-looking* approaches to prediction. Past-looking models assume that the future will resemble the past and therefore extrapolate from recent experience. Future-looking approaches instead posit that the future may diverge sharply from past observations, yet remain predictable if one can infer the stable, underlying generative dynamics. The above distinction closely parallels the Lucas critique in macroeconomics (R. E. Lucas, 1976), which argues that policy evaluations based on low-level historical regularities may fail when the policies themselves would alter the decision environment. Similarly, in asset selection and switching, an agent's own investment decisions reshape the dynamics of future performance. As in the Lucas critique, accurate forecasting thus requires reasoning not from surface-level trends but from the structural parameters that govern asset growth.

Building upon previous work in Bayesian evidence accumulation and function learning, we introduce a normative framework for understanding asset selection and switching decisions. In parallel to Lucas' critique, the framework infers structural parameters that remain stable under interventions, and thus offers a future-looking perspective beyond myopic inferences based on past trends. Across

four behavioral experiments ( $N = 460$ ), we evaluate the proposed forecasting model against myopic heuristic models inspired by the literature on patch foraging and study time allocation. The results highlight forecasting as a core cognitive mechanism for decisions, where long-term consequences emerge endogenously through continued action.

The remainder of the article is organized as follows. We first survey existing frameworks for dynamic decisions, identify their limitations in capturing asset selection and switching problems, and motivate Bayesian function learning as a computational mechanism for human forecasting. We then present formal models that feature either a forecasting component or myopic tendencies, as inspired by past literature. Subsequently, we introduce the four behavioral experiments and compare model fits to human behavior. Finally, we discuss theoretical implications for decision making under endogenous growth dynamics and outline future directions for extending the framework.

## Related Frameworks for Dynamic Decisions

In this section, we introduce existing frameworks for dynamical decisions and contrast their assumptions with those of asset selection/switching. See Table 1 for a summary.

## Past Frameworks and Their Limitations

The first relevant framework is Markov decision processes (MDPs). By defining a transition function that captures environmental dynamics, MDPs have laid the foundation for numerous RL algorithms. This framework has been extensively linked to human decision making and planning (Daw et al., 2011; Dayan & Daw, 2008; Niv, 2009). At the core of many RL algorithms is the Bellman equation, which operates recursively under two key assumptions: (a) the Markov property—state transitions depend only on the current state and action, not the past—and (b) time homogeneity—the probability of transitioning between states under a given action remains invariant over time (Sutton & Barto, 2018). These assumptions become restrictive in dynamic decision problems where asset performance exhibits long-range temporal dependencies. To see this, suppose we have two assets with identical current performance, thus sharing the same state under a standard MDP formulation. However, one asset's performance may have already plateaued, while the other is still improving. The Markovian transition structure of time-homogeneous MDPs cannot distinguish these two cases, making it inadequate for forecasting future asset performance. While this problem could be addressed by encoding all past time points within a single state representation or defining a separate MDP for each asset, these approaches can quickly become computationally intractable and hard to generalize across assets due to the high dimensionality and massive redundancy involved (Sutton & Barto, 2018). While RL algorithms that relax time homogeneity exist (Padakandla et al., 2020), they primarily focus on change point detection rather than modeling long-range temporal dependencies, and their applications to human behavior remain rare.

A second framework for dynamic decisions is patch foraging, which models how agents exploit resources within their current patch before deciding whether to move elsewhere. The normative MVT prescribes leaving a patch when its reward rate falls below the average reward rate across all available patches (Charnov, 1976). A



**Table 1***Comparison Between Asset Selection/Switching and Other Dynamic Decision Problems*

| Modeling framework         | Dynamically evolving alternatives | Maximize cumulative reward over time | Require time-homogeneous transitions | Monotonically decreasing rewards | May return to a previous asset |
|----------------------------|-----------------------------------|--------------------------------------|--------------------------------------|----------------------------------|--------------------------------|
| Asset selection/switching  | Yes                               | No                                   | No                                   | No                               | Sometimes                      |
| Time-homogenous MDP and RL | Yes                               | Yes                                  | Yes                                  | No                               | Yes                            |
| Patch foraging             | Yes                               | Yes                                  | No                                   | Yes                              | Sometimes                      |
| Study time allocation      | Yes                               | Yes                                  | No                                   | No                               | Yes                            |
| Optimal stopping           | No                                | No                                   | N/A                                  | N/A                              | No                             |

*Note.* The columns represent key modeling assumptions. Entries indicate whether a framework typically satisfies the assumption. *Dynamically evolving alternatives* indicates whether option values change over time due to intrinsic dynamics. *Maximize cumulative reward over time* refers to objectives that sum rewards across time steps rather than focusing on performance at a specific time point. *Require time-homogeneous transitions* denotes whether state transitions are stationary and depend only on the current state (Markov property). *Monotonically decreasing rewards* assumes option values always decline over time (e.g., resource depletion). *May return to a previous asset* indicates whether previously selected options can be revisited. Asset selection/switching involves dynamically evolving alternatives, emphasizes performance at a time point over cumulative reward, allows nonstationary transition dynamics, does not assume monotonic decreasing rewards, and may permit returning to a previous asset. These assumptions are typically not satisfied in other frameworks. MDP = Markov decision process; RL = reinforcement learning.

key assumption underlying MVT and its related heuristics is that resource availability in a patch decreases over time with continued consumption (Charnov, 1976; Gibb, 1962; Green, 1984; Hall-McMaster et al., 2021; Kilpatrick et al., 2021; Krebs, 1973; Lin & von Helversen, 2023); resources within the current patch are assumed to decline monotonically. However, many real-world dynamic decisions violate this assumption. In cases where continued investment leads to growth—such as skill development, research progress, or project outcomes—the chosen asset would improve over time and eventually outperform undeveloped alternatives. In these scenarios, foraging algorithms prescribe a myopic strategy of comparing current asset performance, neglecting the forecasting of future performance altogether (Constantino & Daw, 2015). This raises an important question: Under what conditions do foraging-based strategies break down in such settings, and do humans still adhere to them despite the potential drawbacks?

A third relevant domain is study time allocation, which examines how individuals distribute time across multiple learning tasks to maximize their summed final performance at some time horizon. This objective is computationally equivalent to maximizing cumulative performance gains over time, making it more akin to foraging than asset selection/switching (which focuses instead on maximizing the final performance of the single chosen asset). Empirical studies in study time allocation suggest that people prefer tasks that they are worse at and tasks that they can improve faster at (Metcalfe & Kornell, 2005; Son & Kornell, 2008; Tekin, 2022; Ten et al., 2021). Some models incorporate improvement velocity indirectly through a foraging-like mechanism, where continued engagement is reinforced by rewards that decrease monotonically over time (Green, 1984; Payne et al., 2007). Like foraging models, study time allocation models remain inherently myopic, as they do not explicitly forecast future asset performance. In addition, these models assume that the agent can switch back to an old task to build it further, whereas in asset selection and switching problems, it depends on the situation—a researcher may switch back to a previous project, while an employee who quits a position usually cannot return later.

A final framework is optimal stopping, which models decision making in sequential choice tasks where agents must select the best available option upon its presentation, without the ability to return to previously encountered options (Taylor, 1968). The predominant

models for explaining human decisions in these tasks are threshold models, where an option's value is compared against a position-dependent threshold, and selection occurs once the threshold is exceeded (Baumann et al., 2020; Guan et al., 2014; Lee, 2006). Although optimal stopping introduces a temporal component, its structure is fundamentally misaligned with asset selection/switching problems. Here, dynamism exists only in the presentation of options, not in their development over time. Once an option is chosen, its value remains fixed, and there is no subsequent notion of switching. Even if the entire sequence of presented options were treated as a fluctuating time series of utilities, the framework still lacks the key feature of decision making among dynamically evolving alternatives. Given such incompatibility, we will not consider optimal stopping models during our investigation.

## Modeling Background and Motivations

Our Bayesian framework, which integrates extrapolation and comparison, draws inspiration from prior research in human forecasting, function learning, and evidence accumulation. In this subsection, we will outline the key theoretical motivations and connections that inform this approach.

## Human Forecasting and Function Learning

Asset performance trajectories can be naturally framed as time series—one-dimensional functions that evolve over time. From this perspective, agents must infer patterns in an asset's past performance, represent them as functions, and extrapolate these trends to anticipate future outcomes. Importantly, this extrapolation need not rely on simply extending recent performance trends, as criticized in Lucas' account (R. E. Lucas, 1976). In contrast to such surface-level continuation, the Bayesian approach infers latent generative parameters—in our domain, the underlying structural features that govern how performance evolves with continued investment—and combines these inferences with prior beliefs about plausible parameter values to forecast future outcomes (Griffiths et al., 2024; C. G. Lucas et al., 2015; Ma et al., 2023). Here, we review the theoretical foundations supporting this view, drawing from research on human forecasting and function learning.

The human ability to forecast future outcomes has long intrigued researchers from both economic and psychological traditions (Andreassen, 1990; De Bondt, 1993; Harvey et al., 1994; Jones, 1979; Sanders & Ritzman, 1992). While the human forecasting of time series—often referred to as “judgmental forecasting”—demonstrates various biases (Goodwin & Wright, 1993), they frequently match the performance of quantitative forecasting models in real-world applications (Lawrence et al., 2006). Given the ecological relevance of predicting future outcomes, how humans achieve such forecasting accuracy is a problem with both scientific and practical importance.

A significant body of research suggests that humans rely on extrapolation when explicitly forecasting time series, although the precise cognitive mechanisms underlying this process remain debated (Andreassen, 1990; Harvey et al., 1994). The tendency to extrapolate—using past data points to inform future predictions—likely stems from an inductive bias that assumes past trends will continue (De Bondt, 1993). Empirically, this reliance on past data is evident not only in the forecasting of artificially created functions (DeLosh et al., 1997; Harvey et al., 1994) but also when augmented with empirical contexts such as stock market predictions and expert judgments (Hohle & Teigen, 2015; Johnson & Tuckett, 2022). Even in cases where domain-specific knowledge discourages extrapolation, human forecasters often still exhibit this tendency (Johnson & Tuckett, 2022).

Early models of human forecasting commonly assumed linear extrapolation, positing that individuals expect a perceived trend to persist indefinitely (Andreassen, 1990; Harvey et al., 1994). However, subsequent empirical findings challenged this assumption. First, when faced with time series that exhibit stable linear trends, human forecasts do not strictly follow these trends but instead show damping, where predicted growth or decline is attenuated relative to a purely linear extrapolation (Harvey & Reimers, 2013; Lawrence et al., 2006). Second, people do not rely on a fixed extrapolation rule but adjust their forecasts based on the functional form of the underlying trend and the degree to which past data points inform future values (DeLosh et al., 1997; Li et al., 2023). These observations indicate that human forecasting incorporates more than simple linear extrapolation.

A crucial factor in resolving these discrepancies is contextual information. In real-world forecasting, we rarely encounter time series in isolation; rather, we interpret them in light of their semantic context, our prior experience with similar time series, as well as external knowledge (e.g., knowing that an energy shortage could influence stock prices; Harvey & Reimers, 2013; Lawrence et al., 2006). This integration of contextual information is particularly beneficial when the observed time series is noisy and offers little information for forecasting (Sanders & Ritzman, 1992, 1995). Moreover, incorporating context resolves key empirical puzzles regarding damping effects. Specifically, damping can be understood as a regression-to-the-mean effect: When a time series exhibits higher-than-average growth, people tend to damp their predictions downward, whereas for below-average growth, they may antidamp—overestimating future changes (Harvey & Reimers, 2013). Similarly, in environments where past time points are highly informative, human forecasts reflect strong extrapolation; in contrast, when past values are weakly predictive, people shift toward a learned average trend across time series (Li et al., 2023).

Despite progress in identifying the role of context in human forecasting, questions remain on how contextual information is

quantitatively integrated into decision making. A promising approach builds on Bayesian inference, which offers a normative account of how contextual knowledge should be combined with past observations to refine future predictions (Harvey & Reimers, 2013). This perspective is supported by adjacent research on Bayesian function learning, which suggests that humans can indeed learn functions from sequentially presented data points, which mirrors the real-world presentation of time series (C. G. Lucas et al., 2015; Schulz et al., 2017). It has also been shown that Bayesian function learning is used to support downstream decision tasks like multiarmed bandits (Schulz et al., 2018, 2020; Wu et al., 2018). This is similar to the role of function learning in our framework, though without long-term forecasting. Last, recent studies highlight humans’ capacity for metalearning functions (Li et al., 2023), which involves representing and memorizing priors over functions. This metalearning ability is consistent with the empirically observed damping and antidamping tendencies in the forecasting literature and further supports our framework’s assumptions on prior knowledge.

While significant progress has been made in understanding forecasting and function learning, their connection to asset selection and switching remains underexplored. Most quantitative studies focus on functions in isolation, emphasizing estimation and prediction rather than how forecasts guide downstream decision making (Harvey & Reimers, 2013; Li et al., 2023; C. G. Lucas et al., 2015). Some work has linked forecasting mechanisms to value estimation in RL (Schulz et al., 2018, 2020; Wu et al., 2018), but less attention has been given to their role in dynamic decisions with lasting consequences. As a result, it remains unclear how humans integrate forecasting into asset selection and switching or whether such cognitive processes play a direct role in these decisions.

### *Bayesian Evidence Accumulation*

After forecasting each asset’s future performance, the agent must compare these projections in a way that informs asset selection and switching decisions. The decision to stay or switch can be framed as a two-alternative forced-choice problem under noisy evidence inputs, a paradigm extensively studied in cognitive science featuring deep connections with Bayesian inference (Bogacz et al., 2006; Ratcliff & McKoon, 2008). Traditional evidence accumulation models often assume that decision makers track some measure of relative advantage over time. For example, in the classical drift-diffusion model for perceptual decisions, this advantage is represented as the log posterior ratio (i.e., the difference in log posterior probabilities) between two alternatives (Ratcliff & McKoon, 2008). Similar difference computations underpin more recent evidence accumulation models developed for value-based decision making, which have been successful in capturing both empirical choice and decision time data (Fontanesi et al., 2019; Pedersen et al., 2017; van Ravenzwaaij et al., 2020). Our modeling approach follows this principle by assuming that agents track the difference in the posterior or prior mean final performance between options over time, using this as the basis for asset selection and switching decisions.

A limitation of the advantage-based approach is that it is specifically suited for two-alternative forced-choice tasks, where a difference term is computable. This setup captures asset selection between a pair of assets, as well as switching decisions where the agent transitions into a randomly selected new asset. However, in scenarios where the agent must choose among multiple assets or has

the flexibility to select a specific new asset to switch into (Hall-McMaster & Luyckx, 2019), the framework requires extensions to accommodate multialternative evidence accumulation. Addressing this challenge connects to a broader class of models designed for multialternative evidence accumulation (Brown & Heathcote, 2008; Kilpatrick et al., 2021; Tajima et al., 2019). Since the broader domain of asset selection and switching remains relatively understudied, we begin with the simplest two-alternative forced-choice case as a foundation for understanding these decisions. Extending the framework to more complex task setups remains a valuable future direction.

### Theoretical Framework and Computational Models

For asset selection and switching problems, we propose a normative framework that forecasts asset performance trajectories into the future and estimates asset final performance at the time horizon. In the following sections, we will quantitatively formulate the general tasks of asset selection and switching, the normative model, and other alternative models of interest.

### Asset Selection and Switching Task Setups

In both asset selection and switching, the agent's objective is to maximize the final performance of the asset they hold onto at a specified future time horizon. For example, a student might switch between course projects, but can only submit one for evaluation. Similarly, jobs and relationships we leave do not count toward our reward anymore. In some cases, the time horizon is explicitly set, such as a final project deadline. In other cases, the time horizon can be seen as an approximation of when this reward might matter (e.g., an employee's rank at retirement). These deadline-constrained task structures are pervasive across modern professional, educational, and personal domains (Ariely & Wertenbroch, 2002; Rosa, 2013; Son & Kornell, 2008).

The task setup is intentionally simplified to isolate the core computational problem of forecasting future asset performance and making corresponding decisions. We list the main assumptions below:

1. *Agents possess prior knowledge about the typical performance trajectories of different asset types. In asset switching contexts, they also observe the historical performance of their current asset* (how it has evolved with continued investment) and can integrate this information with prior expectations.
2. *Assets improve only when the agent actively invests in them, and remain static otherwise.* Each asset could thus be described by a performance trajectory, specifying how its performance changes over time since the start of its investment. This assumption mirrors real-world cases in which progress arises only through sustained engagement, including employees not yet hired, inactive projects, or relationships not yet pursued. It also implies some degree of mutual exclusivity of investment at any single timestep, common in the same scenarios above. In general, assets tend to improve under sustained investment—some assets, such as course projects, exhibit monotonically increasing performance, while others, like jobs or relationships, may fluctuate. We focus on the increasing case in Experiments 1 to 3, as it is the possibility of improvement that renders past algorithms (e.g., foraging) suboptimal and underscores the need for new frameworks. We address nonmonotonic trajectories in Experiment 4.
3. *The time horizon is constant across assets and does not reset following a switch.* Thus, if the agent switches halfway through, they could only develop the new asset in the remaining time left. While the model could easily be extended to handle asset-specific or uncertain time horizons (by projecting performance to the corresponding evaluation point or integrating over possible time horizon values), we maintain this fixed-horizon setup to preserve interpretive clarity in the current experiments.
4. *When switching to a new asset, the previous asset's performance does not carry over.* This reflects realistic cases in which previous investments yield little direct benefit after starting anew, such as replacing a trainee, changing jobs or projects, or ending and beginning new relationships.
5. *For simplicity, the agent is assumed to make at most one switch before the time horizon.* This constraint simplifies the calculation of expected final performances. Allowing recursive switching (or the option to return to a previous asset) would be relevant for certain real-world scenarios (e.g., picking up an old project) but inapplicable to others (e.g., firing trainees or quitting jobs). Their introduction would require RL-like recursive evaluation of the utilities/expected final performances of continuing for one more timestep versus switching now, both accounting for future continue-switch decisions made using the same policy. Given our article's focus on establishing the forecasting component of human temporally extended decisions, we chose to leave this complication for future modeling.

### The Bayesian Final Performance Model

The Bayesian final performance (BFP) model integrates past performance data with prior knowledge on the general shapes of performance trajectories to make posterior predictions about future performance.

We begin by presenting the model for asset switching, as asset selection is a special case with  $t = 0$ , before any data have been observed. In general, at every timestep, the BFP model forecasts the expected final performances associated with two candidate actions—continuing developing the current asset versus switching to a new asset immediately—at the time horizon. It then uses their forecasted final performance difference to make probabilistic continue-switch decisions. For the current asset, the model integrates its observed performance history with prior expectations on typical trajectories to estimate its final performance. In contrast, a new asset that could be switched into lacks such historical data, so the model incorporates this uncertainty by relying on the expected performance trajectory of a typical asset. We outline the quantitative model below.

Time is indexed by  $t \geq 0$ . The time horizon (i.e., the time point at which asset performance is evaluated) is denoted as  $T > 0$ . For an asset  $i$ , its expected performance trajectory is represented as  $\rho(\tau; \theta_i) \in \mathbb{R}$ , where  $\tau \geq 0$  denotes the relative time since the beginning of asset

investment and  $\theta_i$  denotes the trajectory's parameters. We consider  $\rho(\tau; \theta_i)$  as an intrinsic property of the asset, specifying how its performance improves under sustained investment. The expected performance trajectory may take any arbitrary functional form; our models used monotonically increasing logistic curves, which we define later. The agent aims to maximize the asset final performance at  $t = T$  through asset selection or stay-switch decisions.

We assume that agents have already learned a prior over trajectory parameters,  $p(\theta_i)$ , and that the agent starts with investing in asset  $i$  from time 0 to the current time  $t > 0$ . The current observed performance for asset  $i$ , denoted  $r_i(t) \in \mathbb{R}$ , is assumed to follow a univariate Gaussian distribution centered at  $\rho(\tau = t; \theta_i)$  with a standard deviation  $\nu$ , capturing external noise that affects momentary asset performance (or observation noise):

$$r_i(t) \sim N(\rho(t; \theta_i), \nu^2). \quad (1)$$

Because the agent is currently investing in asset  $i$ , let  $r_i(t) = (r_i(0), r_i(1), \dots, r_i(t))$  represent all performance observations of asset  $i$  up to and including  $t$ . The agent updates its posterior distribution over the trajectory parameters according to Bayes' rule:

$$\begin{aligned} p(\theta_i | r_i(t)) &\propto p(\theta_i) p(r_i(t) | \theta_i) \\ &= p(\theta_i) \prod_{t'=0}^t N(r_i(t'); \rho(t'; \theta_i), \nu^2). \end{aligned} \quad (2)$$

Using the updated posterior, the agent computes the posterior predictive expectation of the current asset's final performance  $r_i(T)$ :

$$\mathbb{E}[r_i(T) | r_i(t)] = \int \rho_i(T; \theta_i) p(\theta_i | r_i(t)) d\theta_i. \quad (3)$$

The agent also calculates the prior predictive expectation of the final performance  $\mathbb{E}[r_j(T)]$  for a randomly sampled new asset  $j$  if they were to switch now at  $t$ . The final performance of the new asset reflects the time remaining to improve it from its baseline, which is  $(T - t)$ :

$$\mathbb{E}[r_j(T)] = \int \rho_j(T - t; \theta_j) p(\theta_j) d\theta_j. \quad (4)$$

The equivalence of priors  $p(\theta_i)$  and  $p(\theta_j)$  depends on the task's generative process. The new asset's performance trajectory could be sampled from the same distribution as the current asset. It could also reflect a new asset type that only became available after  $t = 0$ , featuring a fundamentally different trajectory distribution. The nonequivalent and equivalent scenarios are tested in Experiments 2 versus 3 and 4, respectively. In either case, the agent computes the expected performance difference between staying and switching, which reflects the advantage of the former option over the latter:

$$\Delta = \mathbb{E}[r_i(T) | r_i(t)] - \mathbb{E}[r_j(T)]. \quad (5)$$

To produce a stay-switch decision for the current time  $t$ , the agent feeds the expected performance difference into a logistic decision rule:

$$P(\text{stay}) = \frac{1}{1 + \exp(-k(\Delta - b))}, \quad (6)$$

where  $k$  and  $b$  are the slope and offset parameters of the logistic function, capturing decision noise and bias, respectively.

Asset selection is equivalent to computing  $\Delta$  between two assets  $i$  and  $j$  with different prior distributions  $p_i(\theta_i)$  and  $p_j(\theta_j)$ , in the absence of performance observations at  $t = 0$ . All other aspects of the model remain the same.

The BFP model has three free parameters  $(\nu, k, b)$  that can be fit to human data. We compare two versions of the model, one with  $b$  as a free parameter and another with  $b = 0$ .

## Alternative Models

To assess whether the BFP model can explain human asset selection and switching better than preexisting models, we contrasted it with three alternative models inspired by different literature traditions (Figure 1). We use the same policy (Equation 5) for all the alternative models. As with the BFP models, we fit models with and without a bias term.

### The Bayesian Velocity Model

The Bayesian velocity (BV) model infers the final performance of each option in a way identical to the BFP model. However, inspired by previous studies on training workers (Xiang et al., 2024), we assume instead that agents prefer to see faster performance improvement from the current time  $t$  to the time horizon  $T$ , even if such decisions may lead to low final performance. This is operationalized through a different definition of the advantage term  $\Delta$ :

$$\Delta_{\text{BV}} = \frac{\mathbb{E}[r_i(T) | r_i(t)] - \mathbb{E}[r_i(t) | r_i(t)]}{T - t} - \frac{\mathbb{E}[r_j(T)] - \mathbb{E}[r_j(t)]}{T - t}, \quad (7)$$

where  $\mathbb{E}[r_i(T) | r_i(t)]$  is the posterior mean final performance of trajectory  $\rho$  computed identically to the BFP model, and  $\mathbb{E}[r_i(t) | r_i(t)]$  is its posterior mean current performance, informed by past observations of the current asset's trajectory. Similarly for the new asset  $j$ ,  $\mathbb{E}[r_j(T)]$  and  $\mathbb{E}[r_j(t)]$  are its prior mean final performance and prior mean current performance, respectively. Intuitively, the BV model prefers to invest in assets that predict greater future growth before the time horizon.

The BV model contains the same free parameters as the BFP model:  $(\nu, k, b)$ .

### The Exploitation Heuristic Model

The exploitation heuristic (EH) model is inspired by MVT in foraging (Charnov, 1976). It is a myopic heuristic model that directly compares the currently observed asset performances, such that its decisions do not depend on the future time horizon  $T$ :

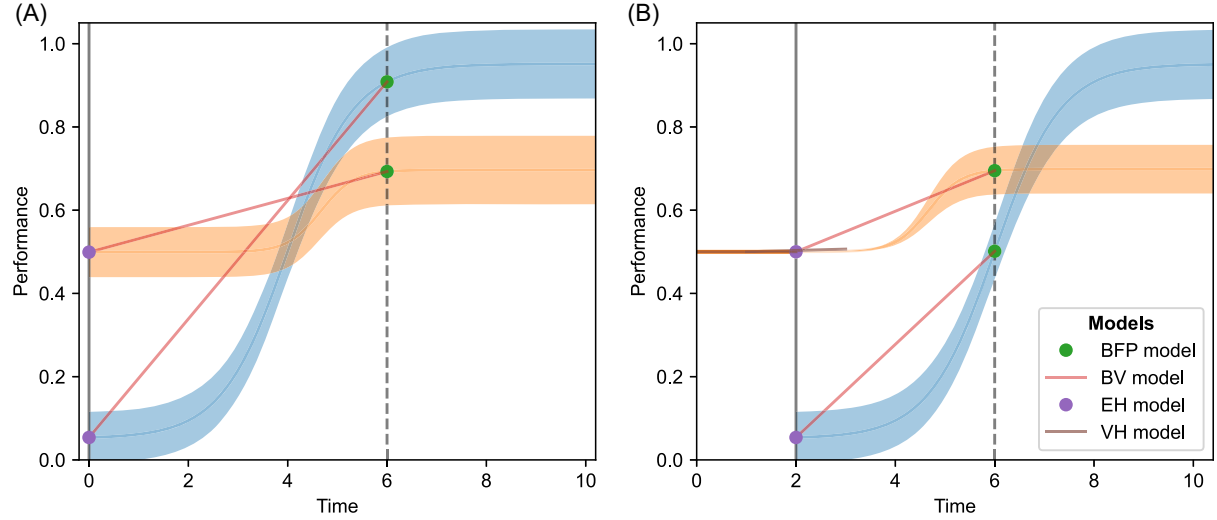
$$\Delta_{\text{EH}} = r_i(t) - \mathbb{E}[r_j(t)], \quad (8)$$

where  $\mathbb{E}[r_j(t)]$  is the new asset's prior mean current performance. This advantage term is akin to MVT, which compares the richness of the current patch with the average patch in the environment to make stay-switch decisions.

Given that the EH model only uses raw observed performances and the prior distribution, there is no Bayesian updating that necessitates the parameter  $\nu$ . Hence, the model only has two free parameters  $(k, b)$ .



**Figure 1**  
Overview of Asset Selection and Switching Models



*Note.* (A) Asset selection. The black solid vertical line denotes current time  $t = 0$ , and the dotted vertical line denotes the task's time horizon ( $T = 6$  used as an example). We assume that agents have some intuitive knowledge of the prior distribution of asset performance trajectories over time (blue and orange denote prior distributions for two example assets; semitransparent lines and ribbons denote prior mean  $\pm$  standard deviation). (B) Asset switching. Agents observe the current asset's performance up to the current time (using the orange asset and time  $t = 2$  as an example) and use this information to estimate the current asset's posterior mean final performance (semitransparent orange lines and ribbons denote posterior mean  $\pm$  standard deviation of the current asset). They also extrapolate how the other asset would develop based on their prior knowledge (semitransparent blue lines and ribbons denote prior mean  $\pm$  standard deviation of the other asset), suppose they were to switch at the current time. For both (A) and (B), the Bayesian final performance (BFP) model compares the prior/posterior mean final performance of different assets at the time horizon (green points), preferring the asset with higher estimated mean final performance. The Bayesian velocity (BV) model computes future velocity (the slopes of red lines) from the current performance of both assets to their prior/posterior mean final performance, preferring the asset with faster growth. The exploitation heuristic (EH) model prefers the asset with higher current performance (purple points). For asset switching, we additionally consider a velocity heuristic (VH) model, which compares the recent velocity of the current asset's performance (between the current time and the previous one) to its average velocity from  $t = 0$  to the current time (the slopes of brown lines; close to flat in this example), preferring to stay if the former is greater. See the online article for the color version of this figure.

### The Velocity Heuristic Model

The velocity heuristic (VH) model is inspired by study time allocation models, which emphasize the role of current performance improvement velocity in determining stay-switch decisions (Metcalf & Kornell, 2005; Son & Kornell, 2008; Tekin, 2022). Notably, it is the only model that does not compare different options to each other, but instead compares the current asset's recent improvement velocity with its average velocity in the past. Hence, the VH model is only possible for asset switching, where the notion of "past velocities" applies when  $t > 0$ . As a myopic heuristic model, it also directly uses raw observations of asset performances to compute recent and average velocities, ignoring the time horizon  $T$ :

$$\Delta_{\text{VH}} = \frac{r_i(t) - r_i(t - \epsilon)}{\epsilon} - \frac{r_i(t) - r_i(0)}{t}, \quad (9)$$

where the first term is the most recent change in velocity over a step size of  $\epsilon$  (which we assume to be 1), and the second term is the average change in observed performance since the beginning of the trajectory. This is akin to study time allocation models that increasingly prefer to switch out of the current asset when its performance plateaus. The VH model has two free parameters ( $k, b$ ).

### Summary of Models

Overall, we have proposed four computational models—the BFP model (three parameters), the BV model (three parameters), the EH model (two parameters), and the VH model (two parameters), each including the logistic offset parameter  $b$ . In addition, we consider the counterpart of each model without the logistic offset parameter  $b$ .

In the context of asset switching, all models incorporate information about the current asset—either its performance history as in the BFP and BV models, its current performance as in the EH model, or its recent and average velocity as in the VH model. The critical distinction lies in the direction of information use. The BFP and BV models are future-looking: They integrate past evidence and infer performance trajectory parameters to forecast what continued investment would yield by the time horizon. The BFP model focuses on forecasting final performance, thus offering normative guidance for the asset selection and switching task setups. The BV model shares this predictive orientation but emphasizes forecasted improvement. Although this focus is misaligned with the task's evaluation criterion, it parallels theories that value growth-oriented progress and intrinsic motivation in human learning (Oudeyer & Kaplan, 2007; Ten et al., 2021; Xiang et al., 2024).

In contrast, the EH and VH models adopt a more past- or present-looking view. Inspired by classical foraging (Charnov, 1976; Green, 1984) and study time allocation frameworks (Metcalf & Kornell,

2005; Son & Kornell, 2008; Tekin, 2022), they rely directly on currently observable outcomes (e.g., the asset's current performance or its recent rate of change) without an explicit forecasting component. These heuristic models thus constitute psychologically plausible baselines informed by previous literature, but remain suboptimal in settings where assets improve at varying rates with continued investment.

In Experiments 2 to 4 (asset switching), we fit all eight models to behavioral data, visualize model predictions using fitted parameter values, and perform quantitative model comparison using fivefold cross-validation. In Experiment 1 (asset selection), the VH model is unable to make predictions. Therefore, we fit the remaining six models and perform model comparison.

## Experiment 1

Imagine an employer deciding which trainee to hire, a PhD student selecting a research direction, or a recent graduate choosing between career paths. Given the temporally extended nature of these decisions, a prudent decision maker must evaluate not only immediate payoffs but also how each asset stands to improve. A trainee may begin with limited skills but learns quickly, a research direction may appear unpromising for a long period before yielding major insights, and a career may involve years of apprenticeship before achievements are recognized or rewarded. Each scenario involves an implicit or explicit deadline—the end of a probationary period, a qualifying exam, a graduation date, or retirement—by which performance is evaluated or simply matters the most. But as of now, the decision maker stands at the beginning, capable of committing to any of these possible futures.

Experiment 1 investigates this class of *asset selection* problems in a simplified but representative setting, exploring how people choose between two candidate assets at time  $t = 0$ . The assets differ in their prior distributions, manifesting as different characteristic shapes. One asset features a higher starting point but slower improvement over time, which we call the “start-high-grow-slow” asset. In contrast, the other asset features a low starting point but faster improvement, which we call the “start-low-grow-fast” asset. The two assets are visualized in Figure 2A. This design elicits qualitatively different model predictions.

To make the experiment intuitive for participants, we asked them to imagine themselves as the leader of a scientific team preparing for an imminent alien attack. They must build a defense system powered by one of two energy types (the “assets” to choose between). Each energy type develops according to a different performance trajectory family (“start-high-grow-slow” or “start-low-grow-fast”), which they observed during training. The participants’ goal was to choose the energy type they would use that would maximize the performance of the defense system at the time of attack.

We asked participants three hypothetical questions, each assuming a different attack time. Critically, participants were not told these time points until after they had observed all the trajectories. This design rules out simple associative strategies in which participants could predetermine fixed mappings between only the queried attack times and asset choices during the training phase (e.g., “if the attack occurs at Month 3, choose Asset A; if at Months 5 and 8, choose Asset B”). Because the attack times were revealed only after the training phase, participants needed to form generalizable representations of asset growth during training in order to perform well in the test phase, so as to forecast asset performance at arbitrary and unanticipated time horizons.

The four models make different predictions. The BFP model predicts that the probability of choosing the “start-low-grow-fast” asset increases with the amount of time left to develop the system. In contrast, the BV model predicts that the highest probability of choosing the “start-low-grow-fast” asset happens at its peak velocity and decreases as the growth asymptotes (anticipating slower future growth). The EH model predicts the same probability of choosing the “start-low-grow-fast” asset across time points. The VH model could not make predictions for this experiment.

## Materials and Method

### Participants

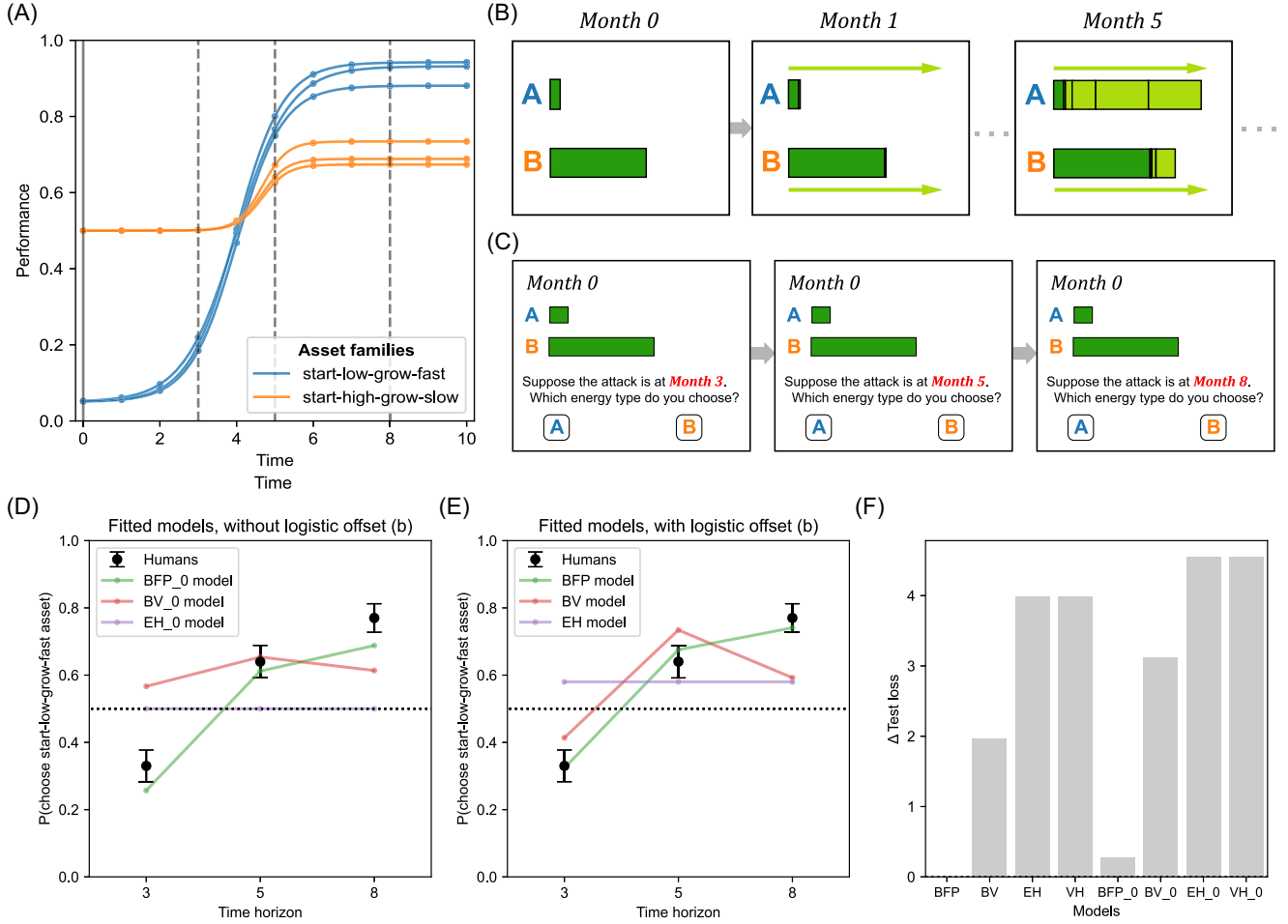
We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. A total of 100 participants (46 women, 53 men, one nonbinary) were recruited on Amazon Mechanical Turk (MTurk). We selected the sample size based on the lowest estimated effect size among dependent variables of interest from a pilot experiment ( $N = 29$ ), to detect the effects with 80% power. Participants received a \$1 base pay and a \$0.5 bonus for making the correct judgment in each scenario. Participants were allowed to participate in the main experiment only after passing a comprehension check following the instructions. We did not exclude any participant or observation. Participants gave informed consent, and the experiment was approved by the Harvard Institutional Review Board (IRB15-2048: Representational Foundations of Adaptive Behavior) and preregistered at <https://aspredicted.org/3jv3-kgxr.pdf>.

### Stimuli

Although prior research in human function learning and graph comprehension demonstrates that people can readily interpret plotted functions (DeLosh et al., 1997; Li et al., 2023; C. G. Lucas et al., 2015), real-world decision contexts rarely provide such explicit visualizations of performance over time. Instead, individuals typically must infer trends by retaining historical performances in memory while observing only their asset's current performance. To better approximate these realistic conditions and avoid the confound that participants might rely on graph-reading skills that do not generalize to real-world settings, we chose not to display performance-time trajectories. Instead, we presented current asset performance as a horizontal progress bar, whose length changes at each queried time point.

For each asset, its horizontal bar starts as a dark green bar showing the starting performance at Month  $t = 0$ , and with each additional month ( $t = 1$  through  $t = 10$ ), a light green bar showing the most recent growth is appended to the right (see Figure 2B for an example). The total length of the bar signifies the system's current performance. The ceiling asset performance of one unit (Figure 2A, vertical axis) corresponded to a total bar length of 1,000 pixels as displayed on participants' devices, which ensured that performance differences between assets were easily discriminable, while keeping the full bar comfortably within the display boundaries of standard laptop and desktop screens.

We modeled performance trajectories using logistic curves, drawing inspiration from prior research demonstrating that logistic growth characterizes many real-world asset trajectories, including human learning, socioeconomic development, population growth, and many more (Harvey & Reimers, 2013; Murre, 2014; Son & Sethi, 2006):

**Figure 2***Experiment 1 Setup, Stimuli, and Behavioral and Modeling Results*

**Note.** (A) The performance trajectories were sampled from two asset families: One starts low and grows fast (blue), and the other starts high and grows slow (orange). We sampled three trajectories from each family. The black solid vertical line denotes the current time (Month  $t = 0$ ). The dotted vertical lines denote the time horizons of each scenario. (B) Participants saw the development of each sampled trajectory. In this example, "A" corresponds to the "start-low-grow-fast" asset and is displayed on top. Each asset's starting performance at Month  $t = 0$  is shown as a dark green segment. Over time, increments in performance are shown as light green segments appended to the right, whose lengths are proportional to the performance improvements. (C) The three test scenarios with time horizons  $T = 3, 5, 8$ . Participants can only observe the starting performance of both assets at Month  $t = 0$  while making selection decisions. (D, E) Behavioral results and model predictions. Black error bars denote the standard error of the mean. The dotted black horizontal line at 0.5 denotes the probability associated with randomly selecting assets. The overlaid colored lines denote the prediction of models using fitted model parameters (D) without the logistic offset  $b$  or (E) with  $b$ . (F) Cross-validation fold-averaged test loss differences. The "0" suffix denotes models fitted without the logistic offset parameter  $b$ . EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance; VH = velocity heuristic. See the online article for the color version of this figure.

$$p(t; \theta_i) = \theta_i^0 + \frac{\theta_i^1}{1 + \exp(\theta_i^2 \cdot (3t - \theta_i^3))}, \quad (10)$$

where  $\theta_i = (\theta_i^0, \theta_i^1, \theta_i^2, \theta_i^3)$ , and each parameter of the logistic curve is sampled from a generative distribution. For the "start-low-grow-fast" trajectory distribution, we have  $\theta_i^0 = 0.05$ ,  $\theta_i^1 \sim N(0.9, 0.03^2)$ ,  $\theta_i^2 \sim N(0.5, 0.03^2)$ , and  $\theta_i^3 \sim N(12, 0.03^2)$ . For the "start-high-grow-slow" distribution, we have  $\theta_i^0 = 0.5$ ,  $\theta_i^1 \sim N(0.2, 0.03^2)$ ,  $\theta_i^2 \sim N(1, 0.03^2)$ , and  $\theta_i^3 \sim N(14, 0.03^2)$ . Participants observed three trajectories sampled from each asset family during training (Figure 2A). These trajectories were the same for every participant, and each green

bar was anchored at the same leftmost position on the screen, extending horizontally to the right.

The performances of the two asset families were displayed one above the other, the order of which was randomized.

### Procedure

In the cover story, participants were instructed to build defense systems in preparation for an upcoming alien attack. They were told that these systems were powered by one of two types of energy (assets), each improving the system's defense capability (asset

performance) at a different rate over time. The participants' goal was to choose defense systems that would maximize defense capability at the designated time of the alien attack (the time horizon). The stimulus display was also explained to participants using example diagrams.

After reading the above instructions, participants completed three comprehension questions assessing their understanding of the task goal, time horizon, and the mapping between total bar length and asset performance. Only those who answered all questions correctly proceeded to the main task. Participants who answered any question incorrectly were asked to review the instructions and repeat the same questions until all were answered correctly, after which they proceeded to the main task. Participants then completed a training block, in which they observed the trajectories of three past defense systems powered by each energy type improve over 10 months, from  $t = 0$  to  $t = 10$ . Participants observed these trajectories sequentially in whatever order they preferred. Each trajectory was shown only once. Participants were never given explicit terminologies such as "start-low-grow-fast" or "start-high-grow-slow," but they may have spontaneously inferred these trajectory distribution properties from the training block.

In the test trials, participants answered three hypothetical questions. Each question asked them to choose which energy type they would use if the attack happened at the end of Month  $T = 3, 5$ , or  $8$ , presented in this order (Figure 2C). Participants responded to each question by making a button click. Importantly, participants could not observe the trajectory development again and had to respond by forecasting how the two systems will develop. They could only see the initial performance of both energy types (dark green bars) on the screen at Month  $t = 0$  when they needed to make the decision.

### Statistical Analysis

We conducted two-sided binomial tests. We report the proportion of participants choosing each energy type and the corresponding  $p$  values according to the binomial tests. We also report the 95% confidence intervals (CIs) constructed using the Clopper–Pearson exact method (Clopper & Pearson, 1934).

### Model Fitting and Comparison

We fit each of the eight candidate models to the binary choice data, assuming a discrete uniform prior over the three performance trajectories observed during training. Since each model had one to three free parameters, they are overparameterized to fit individual participant data (each participant had only three data points). We hence fit each model to the choices generated by all participants, a shortcoming that will be addressed by Experiment 3. We fit model parameters using maximum-likelihood estimation, which connects the model's outputted logistic probabilities to participant choices. Parameter optimization was performed using the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm. We performed the above optimization starting from 10 sets of randomly initialized parameter values, reporting and visualizing results based on the fitted set of parameter values with the lowest loss value.

To compare the models, we performed fivefold cross-validation and juxtaposed the fold-averaged test loss values of each model. This procedure is standard in statistics, machine learning, and

behavioral computational modeling to assess model overfitting (De Rooij & Weeda, 2020; Ghogh & Crowley, 2019; Hastie et al., 2009; James et al., 2013; Kohavi, 1995). Specifically, we randomly partitioned the behavioral data into five folds with equal numbers of data points. We fit each model to every possible combination of four training folds and tested it on the left-out fold (using 10 sets of randomly initialized parameter values). This generated five test loss values, each corresponding to one train–test data set partition. These five values were then averaged as the model's test loss. Overfitted models would predict held-out data poorly and could thus be detected via higher test loss on the left-out folds.

### Data and Code Availability

The online task, data, and analysis code for this and all following experiments are available at [https://github.com/LSZ2001/forecasting\\_guides\\_dynamic\\_decisions](https://github.com/LSZ2001/forecasting_guides_dynamic_decisions) (Liu, 2026).

### Results

Overall, more participants chose the "start-high-grow-slow" asset at  $T = 3$  (proportion = 0.67, two-sided binomial test  $p < .001$ , 95% CI [0.57, 0.76]). In contrast, they prefer the "start-low-grow-fast" asset at both  $T = 5$  (proportion = 0.64,  $p = .007$ , 95% CI [0.54, 0.73]) and  $T = 8$  (proportion = 0.77,  $p < .001$ , 95% CI [0.68, 0.85]). Quantitatively, as the time horizon increases, the proportion of participants choosing the "start-low-grow-fast" asset increased. We fit the three models that could make predictions for asset selection—BFP, BV, and EH—to the behavioral data, with and without the logistic offset parameter  $b$ . As shown in Figure 2D and 2E, the BFP model captures the behavioral pattern that the "start-low-grow-fast" asset is increasingly preferred as the time horizon increases. In contrast, the BV model predicted that the "start-low-grow-fast" asset was most likely to be chosen at the middle  $T = 5$  condition, whereas the EH model predicted the same choice probabilities across all three time horizons. Model comparison via fivefold cross-validation supports the qualitative conclusions above: The BFP models with and without logistic offset  $b$  had the lowest fold-averaged test loss across all candidate models, with including  $b$  providing a slightly better fit (Figure 2F). Model recovery analysis revealed that fivefold cross-validation was capable of identifying the correct generative model, except degenerate cases where fitted models produce near-identical predictions across trials and thus could be approximated by any model with sufficiently large decision noise and an appropriate offset. These models clearly could not capture the empirical human behavior observed, which exhibits systematic, structural variations across trials.

### Discussion

In Experiment 1, we found that participants were sensitive to asset performance trajectories into the future as well as to differing time horizons. Their decisions were best described by the BFP model, compared to the alternative models that use myopic strategies common in the fields of foraging and study time allocation.

### Experiment 2

In reality, decision makers are rarely bound indefinitely to their initial commitments. Employers dissatisfied with a trainee's progress may



consider hiring anew; PhD students disillusioned with their current research direction may consider shifting to a new field; professionals may discover new career paths that did not exist when they first entered the job market. Such decisions to switch inevitably incur a temporary setback, as previous progress cannot be carried over. In practice, this may take the form of lost productivity, slowed research output, or reduced job status, freedom, or visibility. However, the ultimate deadline—evaluation, graduation, or retirement—remains fixed. Consequently, decision makers must carefully weigh the future gains of switching against those of developing their current asset further, despite its imperfect prospects. Choosing whether to persist or to start anew thus poses a distinctive and cognitively demanding decision problem.

Experiment 2 mirrors this class of *asset switching* problems, as we probe participants to make decisions after developing some given asset for a while. Staying with the current asset builds on the current progress, whereas switching to another asset builds from its starting performance. Therefore, participants need to mentally translate the trajectories in order to account for the progress already made. This additional difficulty makes it informative to study asset switching as an independent problem from asset selection and to explore whether humans resort to heuristics instead.

For consistency, we used the same two assets as in Experiment 1. The only difference was that participants were forced to develop a randomly assigned starting asset up to  $t = 2$ , at which point they decide whether to stay with the current asset or switch to another asset for different hypothetical scenarios (Figure 3).

The BFP model chooses to stay with or switch to the “start-high-grow-slow” asset when there is little time remaining to develop the system, and stay with or switch to the “start-low-grow-fast” asset when there is more time. In contrast, all three alternative models prefer the same asset across time horizons.

## Materials and Method

### Participants

A total of 200 participants (103 women, 93 men, one nonbinary, three prefer not to say) were recruited on MTurk. One hundred of them (55 women, 43 men, one nonbinary, one prefer not to say) were assigned to start with the “start-low-grow-fast” asset, and the other 100 participants started with the “start-high-grow-slow” asset. We selected the sample size based on the lowest estimated effect size among dependent variables of interest, according to a pilot experiment ( $N = 50$  per condition) to detect the effects with 80% power. Participants received a \$1.5 base pay and a \$0.5 bonus for making the correct judgment in each scenario. Participants were allowed to participate in the main experiment only after passing a comprehension check following the instructions. We did not exclude any participant or observation. Participants gave informed consent, and the experiment was approved by the Harvard Institutional Review Board and preregistered at <https://aspredicted.org/f33z-s27j.pdf>.

### Stimuli

We used the same stimuli and trajectory distributions as in Experiment 1 (the cover story on defense systems and energy types remains similar), except that participants observed 2 months of development on one randomly chosen energy type before they made decisions to stay or switch in three hypothetical scenarios (Figure 3A

and 3B). In the test trials, the asset they started was always displayed above the other asset (Figure 3C).

### Procedure

After reading the instructions, participants completed four comprehension questions to verify their understanding of the task goal, the correspondence between total bar length and asset performance, the consequence of switching (i.e., developing a new asset from its initial performance within the remaining timesteps), and the rule that they would be evaluated based on the asset they were developing at the time horizon. Only participants who answered all questions correctly proceeded to the main task.

Participants then completed two training blocks. The first training block was the same as Experiment 1, where they observed the trajectories of three past defense systems powered by each energy type improve from  $t = 0$  to  $t = 10$ . Participants observed each trajectory sequentially in whatever order they preferred. Each trajectory was shown only once. In the second training block, they first observed a defense system develop with one energy type for 2 months and then observed how the performance changes with either continuing with the same energy type or switching to the other energy type, from  $t = 2$  to  $t = 10$ . Participants observed each trajectory sequentially in an interleaved order, where an example of staying with the current energy type was followed by an example of switching to the other energy type, and vice versa. Each trajectory was shown only once.

In the test trials, participants answered three hypothetical questions. Each question asked them whether to continue with the current energy type or switch to the other energy type if the attack happened at the end of Month  $T = 3, 5$ , or  $8$ , presented in this order (Figure 3C). Participants responded to each question by making a button click. Importantly, participants could not observe the trajectory development again and had to respond by forecasting how the two systems will develop. They could only see the first 2 months of performance of the current energy type and the initial performance of the other energy type on the screen at Month  $t = 2$  when they needed to make the decision.

### Statistical Analysis

The statistical analysis was identical to Experiment 1.

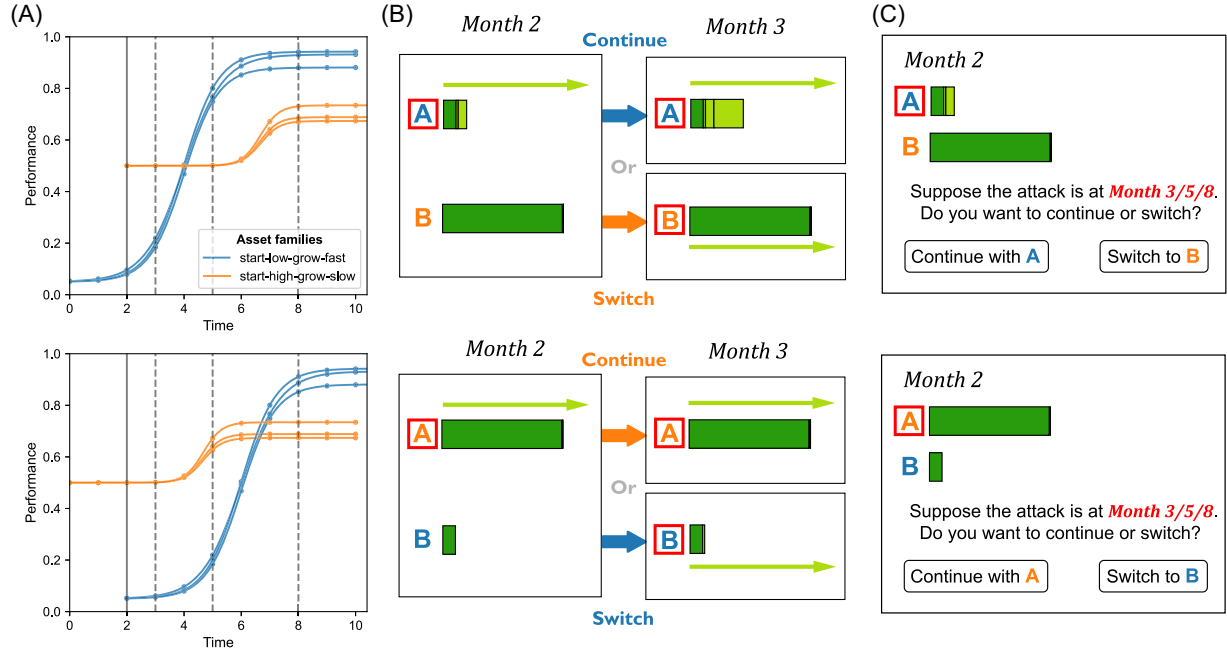
### Model Fitting and Comparison

The model fitting and model comparison procedures were identical to Experiment 1.

## Results

Overall, among participants who were assigned the “start-low-grow-fast” asset to begin with, more participants chose to switch when  $T = 3$  (proportion = 0.73,  $p < .001$ , 95% CI [0.63, 0.81]), whereas more participants chose to stay when  $T = 5$  (proportion = 0.67,  $p < .001$ , 95% CI [0.57, 0.76]) and  $T = 8$  (proportion = 0.83,  $p < .001$ , 95% CI [0.74, 0.90]). Among participants who were assigned the “start-high-grow-slow” asset to begin with, more participants chose to stay when  $T = 3$  (proportion = 0.83,  $p < .001$ , 95% CI [0.74, 0.90]) and  $T = 5$  (proportion = 0.62,  $p = .021$ , 95% CI [0.52, 0.72]), whereas

**Figure 3**  
Experiment 2 Setup, Performance Trajectories, and Display



**Note.** Different rows correspond to the different between-participant task conditions. (A) The three performance trajectories sampled for each trajectory distribution. Trajectory colors denote the different asset families. The black solid vertical line denotes the current timestep (Month  $t = 0$ ), and the dotted vertical lines denote the time horizons of different test scenarios. (B) Visualization of asset performance over time. Participants must develop their task condition's assigned starting asset up to Month 2, from which they can decide to continue or switch to the other asset, developing it in the remaining time before the time horizon. (C) The three test scenarios with  $T = 3, 5, 8$  as presented to participants. They can only see the starting performance of the new asset, as well as the current performance of their starting asset at the designed decision time of Month  $t = 2$ . See the online article for the color version of this figure.

more participants chose to switch when  $T = 8$  (proportion = 0.9,  $p < .001$ , 95% CI [0.82, 0.95]). These qualitative behavioral patterns were only predicted by the BFP model, whether it was fit with or without the logistic offset parameter  $b$  (Figure 4A and 4B). Model comparison via fivefold cross-validation also shows that the BFP models provide the best fit to the data, with including  $b$  providing a slightly better fit (Figure 4C), while model recovery demonstrated the recoverability of true generative models using this pipeline.

Despite the qualitative match, humans quantitatively deviate from the BFP model. The BFP model predicts that the probability of staying should be near-identical between  $T = 5$  and  $T = 8$  when participants start with the “start-low-grow-fast” asset, and near-identical between  $T = 3$  and  $T = 5$  for the other task condition. This is because the advantage term  $\Delta$  is near-identical for the two test scenarios under the same  $\nu$  value, such that the downstream logistic decision rule could not map them to significantly different choice probabilities. However, among participants, we consistently see higher uncertainty in the  $T = 5$  test scenario, where the choice probabilities are closer to the chance level of 0.5. The other models also fail to capture this phenomenon.

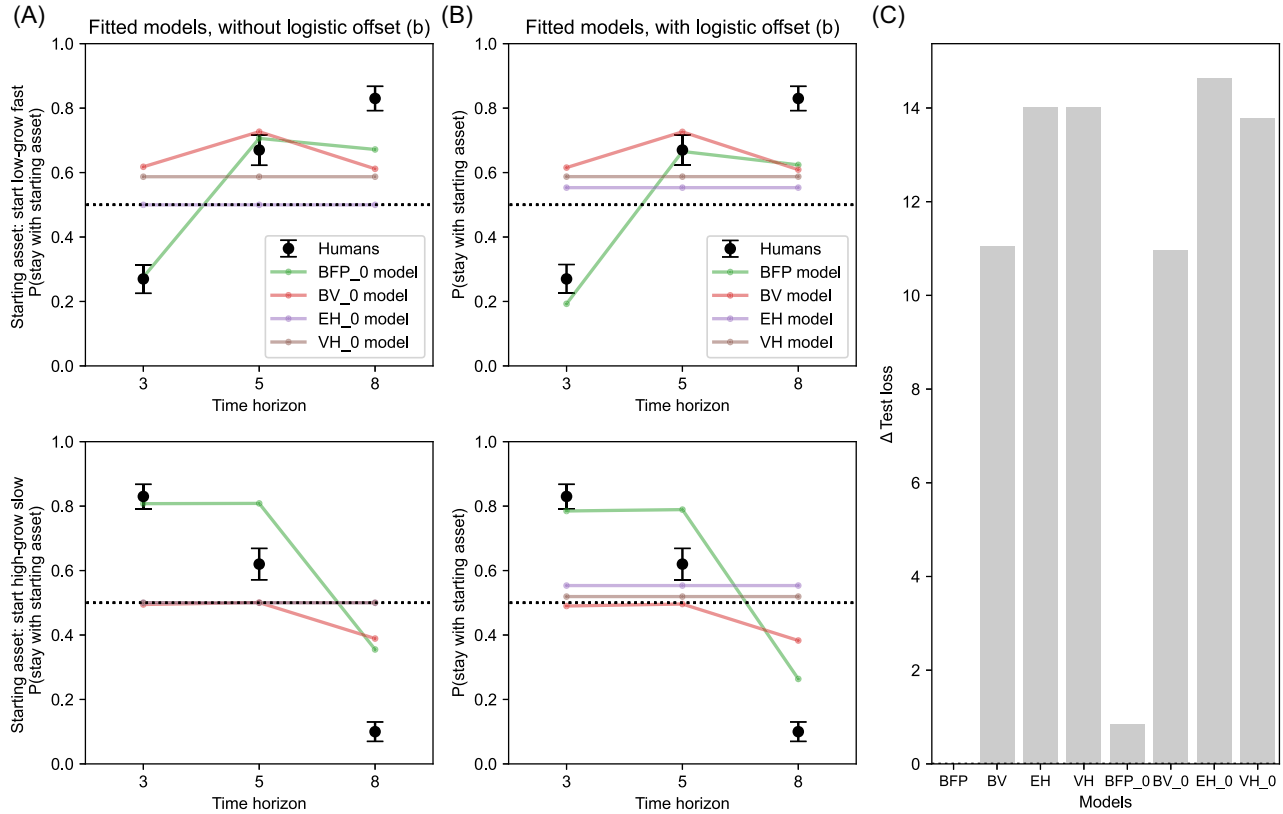
## Discussion

In Experiment 2, we found that participants were capable of extrapolating asset performance into the future even when accounting for prior investment in one asset, and they could use extrapolation and

translation to make decisions that consider the relevant time horizons. Their decisions were best described by the BFP model, compared to the alternative models that use myopic strategies common in the fields of foraging and study time allocation.

It is also worth noting that this task required storing trajectory distributions in working memory and mentally translating the trajectories, which was a considerable challenge that might have swayed participants in the direction of using myopic heuristics instead. Even so, participants continued forecasting future performances and making decisions qualitatively consistent with the BFP model. It is also evident that humans did not misapply the BFP model under the premise of asset selection, which seemed tempting as it would also remove the burden of mental translations. If such misapplication had occurred, participants would consistently prefer the “start-low-grow-fast” asset at  $T = 5$  across both conditions (as in Experiment 1). In contrast, our participants preferred different assets depending on which asset they were assigned to start with.

The unexplained variability of human responses in the  $T = 5$  scenario highlights areas for future model development. Importantly, this variability runs counter to a purely associative explanation arguing that participants memorize fixed mappings between each time point and an optimal asset choice. A tabular model of this kind would predict consistent certainty across timesteps, as opposed to greater uncertainty at  $T = 5$ . Furthermore, such an associative model would still require a forecasting mechanism, because no two trajectories

**Figure 4***Experiment 2 Human Behavior and Modeling Results*

**Note.** (A, B) Behavioral results and model predictions. Black error bars denote the standard error of the mean. The dotted black horizontal line at 0.5 denotes the probability associated with randomly choosing assets. The overlaid colored lines denote the prediction of models using fitted model parameters (A) without the logistic offset  $b$  or (B) with  $b$ . Top row corresponds to the condition where participants started with the “start-low-grow-fast” asset. Bottom row corresponds to the condition where participants started with the “start-high-grow-slow” asset. (C) Cross-validation fold-averaged test loss differences. The “0” suffix denotes models fitted without the logistic offset parameter  $b$ . VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

were presented simultaneously during training—meaning that the growth pattern of one asset family had to be maintained in memory to derive any responses being cached. One possible explanation for the greater response variability at  $T = 5$  arises from the experiment’s sequential presentation format—asset performance is revealed to participants chronologically, a setup typical in real life. Consequently, participants may have retained better memory of the starting and final performance of assets under the well-known primacy and recency effects (Glanzer & Cunitz, 1966), leading to more uncertain decisions at the middle  $T = 5$  compared to  $T = 3$  and  $T = 8$ .

Could it be possible that participants engaged in a weaker form of forecasting, relying on simple linear extrapolation from recent performance trends rather than learning the characteristic trajectories of asset development (Andreassen, 1990; DeLosh et al., 1997; Harvey et al., 1994)? This alternative account cannot explain several aspects of the observed behavior. First, participants had to forecast the performance of new asset families that they could switch into, for which no historical data were available. This absence of past information precludes linear extrapolation based on past trends. Second, when participants began with the “start-low-grow-fast” asset family and made decisions at timestep  $t = 2$ , the observed

performance increments were minimal (typical of the initial flat stage of a logistic growth trajectory). If participants linearly extrapolated solely from the recent increments, they would predict that the “start-low-grow-fast” family would never surpass the initial performance of the “start-high-grow-slow” family, and would thus always prefer switching out of the former. Contrary to this prediction, empirical participant behavior reflected a flexible sensitivity to the accelerating growth pattern characteristic of logistic performance trajectories.

A related possibility is that participants employed a simpler interpolation heuristic, using the asset performances at  $t = 2$  and  $t = 10$  as relatively stable “pivot points” (where memory noise is lowest) and linearly interpolating intermediate values. However, unlike the linear extrapolation account, this interpolation hypothesis must explain how participants obtained an estimate of performance at  $t = 10$  a priori in order to interpolate. Such use of future pivot points necessitates an independent forecasting process akin to that proposed in the BFP model. Still, the above investigation raises an intriguing direction for future research: studying how the time duration from the current time  $t$  to time horizon  $T$  affects human stay-switch strategies and whether longer time durations elicit usage of increasingly myopic heuristics.

### Experiment 3

Experiments 1 and 2 demonstrated that the BFP outperformed the alternative models in both asset selection and switching. However, one limitation was that the experiment designs were simplistic and contrived—for example, the asset trajectory families  $p(\theta)$  were highly distinguishable from each other and rather homogenous within each family, there were only two possible assets at a time, and we enforce a strict decision time  $t = 2$  upon participants in Experiment 2. In real-world situations, these decisions are often more complicated and less tractable. For example, switching jobs usually entails more uncertainty about what it would feel like in a new position, and there is usually not a fixed time point to make these decisions. The enforcement of decision times also prevents us from comparing the BFP model to other popular foraging heuristics, such as the fixed-time and fixed-number patch-leaving strategies (Gibb, 1962; Krebs, 1973; Lin & von Helversen, 2023; McNair, 1982), which assume that individuals flexibly determine their switch times. In addition, given that each experiment only asked for three binary judgments, the models were fitted on all participants' data, which overlooks individual differences.

In Experiment 3, we retain our focus on asset switching but extend the paradigm of Experiment 2 by addressing its earlier limitations. Specifically, we sample the current and new asset performance trajectories from the same distribution featuring large variance and allow participants to switch whenever they want. Each participant completed two conditions with different trajectory distributions—one with trajectories separating closer to the time horizon  $T = 10$  (“separate-late” condition) and the other with trajectories separating closer to the beginning  $t = 0$  (“separate-early” condition; Figure 5A and 5B). In each condition, participants completed 50 runs each lasting from  $t = 0$  to

$T = 10$ , where they must choose to stay or switch at every time point (but they could only switch at most once per run). This design more closely examines the processes underlying participants' decisions and provides more data points that enable model fitting to individual participants.

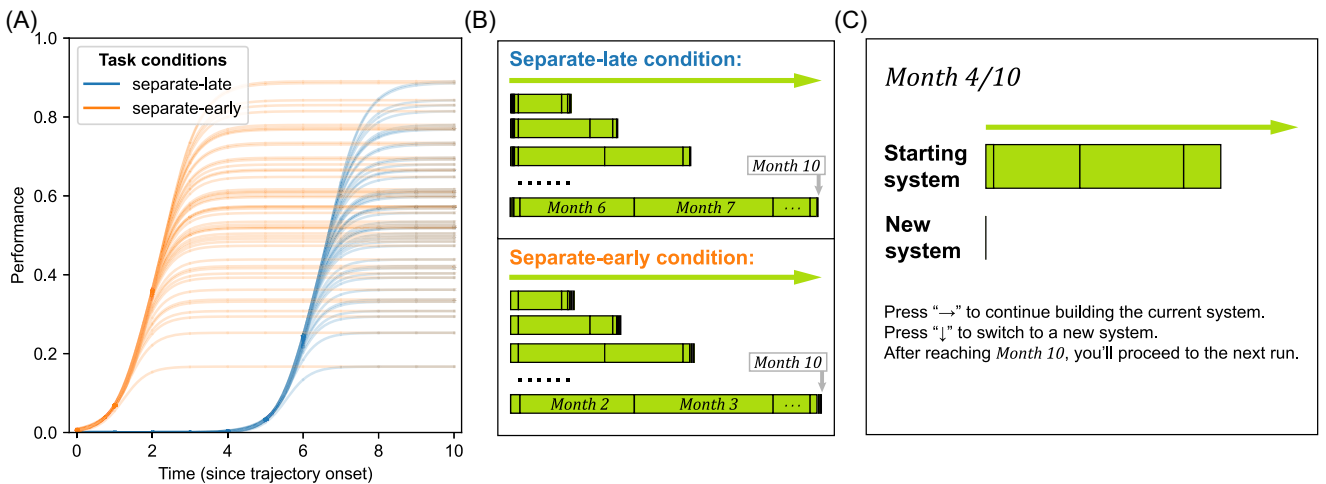
The BFP model makes the following predictions: (a) The mean switch time is earlier in the “separate-early” condition; (b) the mean asset performance at the switch point is higher at the “separate-early” condition; (c) switching is less likely in the “separate-early” condition; (d) comparing runs in which participants stayed until the end versus switched halfway through, the mean final performance of the starting asset should be higher in the former group of runs. In contrast, none of the alternative models makes all four of these predictions. We also compare participants and models in their probability of staying at every time point conditioned on them not switching so far, analyzed both separately and jointly across the two conditions. Finally, we perform fivefold cross-validation to compare the four models (with and without logistic offset  $b$ ).

## Materials and Method

### Participants

A total of 80 participants (36 women, 43 men, one nonbinary) were recruited on MTurk. We selected the sample size based on a pilot experiment ( $N = 29$ ) to detect the effects of interest with 80% power. Participants received a \$6 base pay and a bonus up to \$2 proportional to the mean final performance achieved across test runs. Participants were allowed to participate in the main experiment only after passing a comprehension check following the instructions. We did not exclude any participant or observation. Participants gave informed consent, and

**Figure 5**  
*Experiment 3 Setup, Performance Trajectories, and Display*



*Note.* (A) Example performance trajectories sampled from the “separate-late” (blue) and “separate-early” (orange) conditions. (B) The display of performance trajectories during training. The green bars illustrate how the 50 training trajectories were displayed—rows correspond to different assets sampled from the same trajectory distribution, and they are ordered from lowest to highest final performance. The plateauing region of trajectories corresponds to black smudges at the beginning or end of the green bars, representing light green segments with near-zero width. Subjects could press keyboard keys to toggle back and forth in time to see how all these training trajectories grow, until they felt ready to begin the task. (C) The display of a test run. At each time from  $t = 1$  to  $t = 9$  inclusive, participants must decide to either continue investing in their current defense system/asset or switch to a new one and develop it in the remaining time before  $T = 10$ . The starting and new asset trajectories are both sampled from the condition-specific trajectory distributions visualized in (A). See the online article for the color version of this figure.



the experiment was approved by the Harvard Institutional Review Board and preregistered at <https://aspredicted.org/sh73-d3z4.pdf>.

### Stimuli

The same cover story of defense systems as assets was used. The trajectory distributions differed across the two conditions:

$$\rho(t; \theta_i) = \theta_i^0 + \frac{\theta_i^1}{1 + \exp(\theta_i^2 \cdot (3t - \theta_i^3))}, \quad (11)$$

where  $\theta_i^0 = 0$ ,  $\theta_i^1 \sim N(0.55, 0.15^2)$ , and  $\theta_i^2 = -\frac{1}{2}\theta_i^1 + \frac{9}{8}$  for both conditions. For the “separate-early” condition,  $\theta_i^3 = 17 - \frac{25}{24} + \frac{25}{6}\theta_i^1$ ; for the “separate-late” condition,  $\theta_i^3 = 3 + \frac{25}{6}\theta_i^1$ . In other words, the trajectory distributions  $p(\theta)$  used for the two conditions were horizontally translated versions of each other, where stochasticity was introduced by the single trajectory parameter  $\theta_i = \theta_i^1$  (Figure 5A). Our experiment directly visualized the above  $\rho(t; \theta_i)$ , leading to the simplification  $r_i^{(t)} = \rho(t; \theta_i)$  in the true generative process. Unlike Experiments 1 and 2, the trajectories used in training and test time were randomly sampled for each participant in each run. The starting asset is always displayed above the other asset that could be switched into (Figure 5C).

Due to the greater number of stay-switch decisions participants must make, we changed the response format from clicking buttons to pressing keyboard keys—the “right” arrow to continue developing the current asset and the “down” arrow to switch to a new asset.

### Procedure

After reading the instructions, participants answered four comprehension questions testing for the same understanding as in Experiment 2. Only participants who answered all questions correctly proceeded to the main task. In the instructions, participants were only told that asset performances typically grow, and had to infer condition-specific trajectory distribution properties from the training block examples.

Participants then completed the “separate-late” and “separate-early” task conditions in randomized order. In each condition, participants first observed 50 assets improve from  $t = 0$  to  $t = 10$  simultaneously, where the asset performance trajectories were sampled from the condition-specific prior  $p_{\text{cond}}(\theta)$ . Due to experimental time constraints, the 50 trajectories were presented simultaneously on the screen, ordered from lowest to highest final performance (Figure 5B). During observation, participants could toggle back and forth between time-steps by pressing the left and right arrow keys on their keyboard, and the display of all 50 trajectories would be updated to reflect asset performances at the current timestep. Once the participant felt ready, they could voluntarily end observation and proceed. Importantly, participants could only proceed after toggling to the  $t = 10$  timestep, which ensured that all participants had observed the full growth of the training trajectories from  $t = 0$  to  $t = 10$  at least once.

Participants then completed a second training block consisting of five runs. These five training runs were identical to those in the test block, but included numeric feedback on the final performance achieved after the run ended at  $T = 10$ .

In the subsequent test block, participants completed 50 runs for that condition. In each run, participants must first develop the randomly assigned starting asset for 1 month (till Month  $t = 1$ )—so that there was an asset to stay with or switch out of—but they were

allowed to make stay-switch decisions afterward. If at any point participants decided to switch, the new asset’s performance trajectory would be randomly sampled from the same condition-specific trajectory distribution  $p_{\text{cond}}(\theta)$ . Participants were only allowed to switch once per run to prevent recursion in estimating the utility of switching in all our models. Given this restriction, we did not analyze participants’ stay-switch decisions after they first switched in each run, because from that point on, they were not allowed to switch again, rather than choosing not to switch. Participants were told that their bonus was proportional to the sum of the final performances achieved across all 100 test runs across both conditions.

### Statistical Analyses

We conducted two-sided paired  $t$  tests. In the main text, we report the  $t$ -statistics, their corresponding  $p$  values, the Cohen’s  $d$  effect sizes, and 95% CIs of the effect sizes. The BFP models’ first two predictions were assessed based on all participants who switched at least once in both conditions ( $N = 63$ ), whereas the fourth prediction was assessed based on all participants who switched at least once in either of the conditions ( $N = 79$ ). Prediction 3 was assessed based on all participants ( $N = 80$ ).

### Model Fitting and Comparison

For each participant, we fitted all four models, with and without the logistic offset  $b$ , to their binary stay-switch data, jointly across both conditions. The trajectory prior used was the true generative prior  $p_{\text{cond}}(\theta = \theta^1)$  for each condition, enumerated over a dense grid of  $\theta^1$  values for numeric integration because the logistic function families prevent analytic solutions for integrals over parameter values. Similar to Experiments 1 and 2, we fit model parameters using maximum-likelihood estimation, which connected the model’s outputted logistic probabilities to participant choices. Due to the greater data set sizes in Experiment 3, optimization was performed by fitting all model parameters using the PyBADS optimizer (Acerbi & Ma, 2017; Singh & Acerbi, 2023). We performed the above optimization starting from 10 sets of randomly initialized parameter values, reporting and visualizing results based on the fitted set of parameter values with the lowest loss value.

To compare candidate models, we again performed a fivefold cross-validation and juxtaposed the fold-averaged test loss values associated with each model. In this process, we randomly partitioned each participant’s data into five folds with equal numbers of data points. For each model fitted on each participant, we optimize it over every possible combination of four training folds and tested it on the left-out fold (also using 10 sets of randomly initialized parameter values), generating five test loss values that each corresponded to one train-test data set partition. These five values were then averaged as the model’s test loss. In the main text, we visualized the fold-averaged test loss summed across all participants to capture each model’s performance in accounting for the behavior of all participants.

### Results

Across both task conditions, the mean number of analyzable stay-switch responses provided by participants was  $701 \pm 13.9$ , excluding the  $t = 0$  and  $t = 10$  times of every run and all the time points within a run after switching because participants could not choose during

these times. Among the 80 participants, 79 participants switched at least once throughout the experiment, and 63 participants switched at least once in both conditions. The number of data points per participant and the nonuniformity of their responses enabled the fitting of participant-specific models and our subsequent interpretation.

Consistent with the BFP model predictions, the mean switch time of participants was later in the “separate-early” condition,  $t(62) = 2.47, p = .016, d = -0.370, 95\% \text{ CI} [-0.689, -0.0687]$  (Figure 6A). The mean performance level at which participants switch was also higher for this condition,  $t(62) = 14.5, p < .001, d = 2.41, 95\% \text{ CI} [-3.01, -1.91]$  (Figure 6B), and they were less likely to stay with their starting asset till  $T = 10, t(79) = 6.52, p < .001, d = 0.719, 95\% \text{ CI} [0.483, 0.981]$  (Figure 6C). Most importantly, participant responses reflected considerations for the starting asset’s final performance—which is also the basis for the BFP model’s judgments. Among the runs they stayed until the end, the starting asset’s final performance was higher than that among runs in which they switched,  $t(78) = 10.9, p < .001, d = -1.90, 95\% \text{ CI} [-2.40, -1.47]$  (Figure 6D).

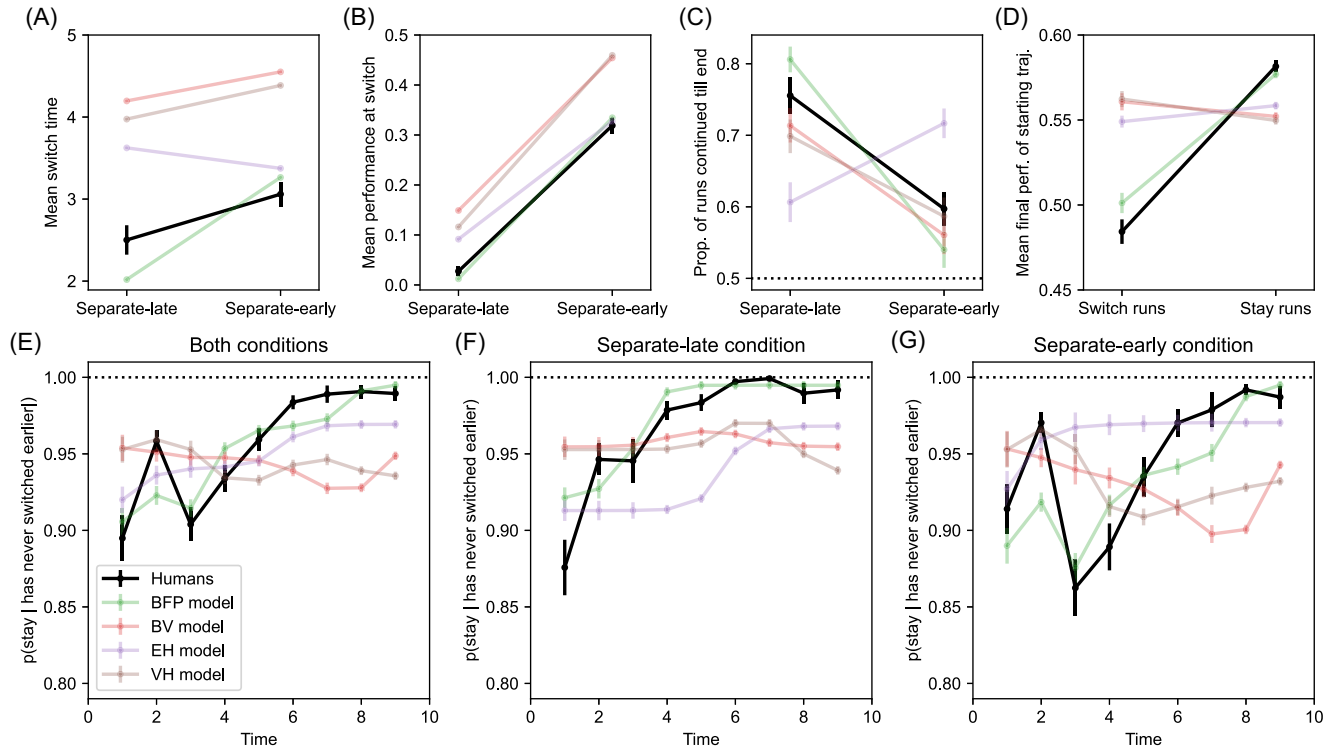
We compared the above behavioral patterns to the predictions generated from all four models with and without logistic offset  $b$  fitted to the behavioral data. The models fitted with  $b$  are visualized in Figure 6; for models fitted without  $b$ , see Figure 7. Only the BFP

model with logistic offset could quantitatively capture the behavioral patterns across conditions and stay versus switch runs.

In addition, we compared the data and model predictions in terms of the probability of staying at each time point, conditional on the agent not switching yet—for data pooled across conditions (Figure 6E) and for each condition (Figure 6F and 6G). This analysis was based on the actual runs each participant completed. The BFP model with logistic offset captured the qualitative and quantitative patterns of the behavioral data, including the decreasing tendency to switch over time in the “separate-late” condition, as well as the surge in switching probability around  $t = 3$  in the “separate-early” conditions when the trajectories had separated. In contrast, all alternative models with fitted logistic offsets failed to capture these behavioral trends. Additionally, all models fitted without the logistic offset consistently predict chance levels of switching at early timesteps, failing to capture the human tendency to stay (Figure 7E–7G).

Model comparison via fivefold cross-validation similarly favored the BFP model with logistic offset, which has the lowest test loss among all models (Figure 8A). This conclusion is not the result of a few outliers and holds for individual participants: Post hoc two-sided paired  $t$  tests between the BFP model with logistic offset  $b$  and the other models with or without  $b$  are all significant at the  $\alpha = .0001$

**Figure 6**  
*Experiment 3 Human Behavior and Modeling Results*



**Note.** The models visualized have the logistic offset  $b$  as a free parameter. Error bars denote the standard error of the mean across participants. (A–C) Data (black) and model-predicted (color) mean switch time (A), mean starting asset performance at switch (B), and probability of continuing with the starting asset till the end of the run (C), for the “separate-late” and “separate-early” task conditions. (D) The starting asset’s true final performance (if the participant had continued till the end of the run), for runs where participants continued till the end versus switched midway. (E–G) Data (black) and model-predicted (color) probability of staying at each time point given that they have not switched so far in the run, visualized for both task conditions together (E), the “separate-late” condition (F), and “separate-early” condition (G). Model predictions with logistic offset  $b$  are shown. VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

level (Figure 8B and 8C). We also assessed whether a fivefold cross-validation could recover the true generative model through model recovery, by performing the same model-fitting procedures over the predictive data generated by each participant-fitted model. The model recovery results demonstrate that our procedure could reliably recover the BFP, BV, EH, and VH models from each other, but is less powerful in recovering the models without  $b$  (Figure B1). Despite this, model recovery confirms that the BFP model only has the lowest test loss on data it generated (i.e., when it is the true generative model), but not others.

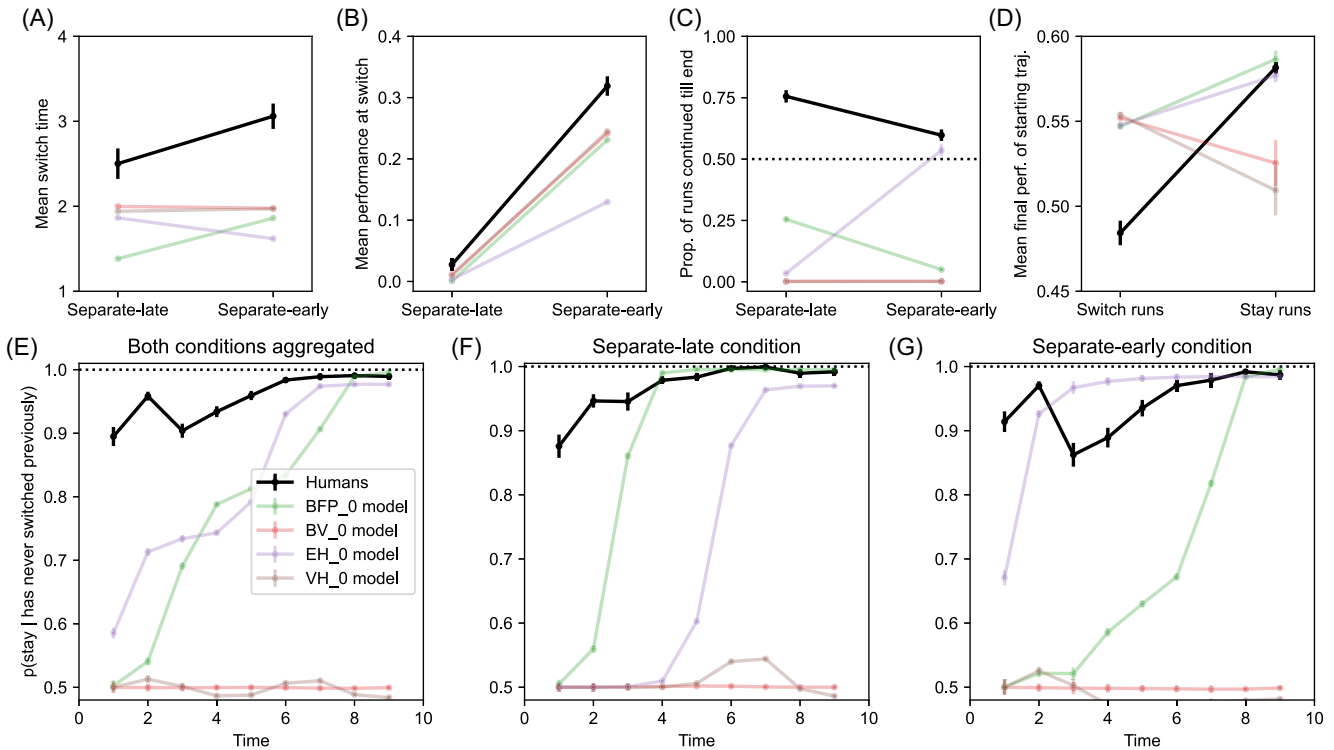
## Discussion

In Experiment 3, we adopted a more naturalistic task setup for asset switching, where the starting and new assets were sampled from the same trajectory distribution and participants were allowed to switch at any time. By fitting the models to each individual participant's data, we showed that the BFP model with logistic offset captured participants' responses and outperformed the other models. Some key features of the behavioral data—such as the mean switch time and mean performance at switch point—could not be explained by the fixed-time and fixed-number foraging heuristics.

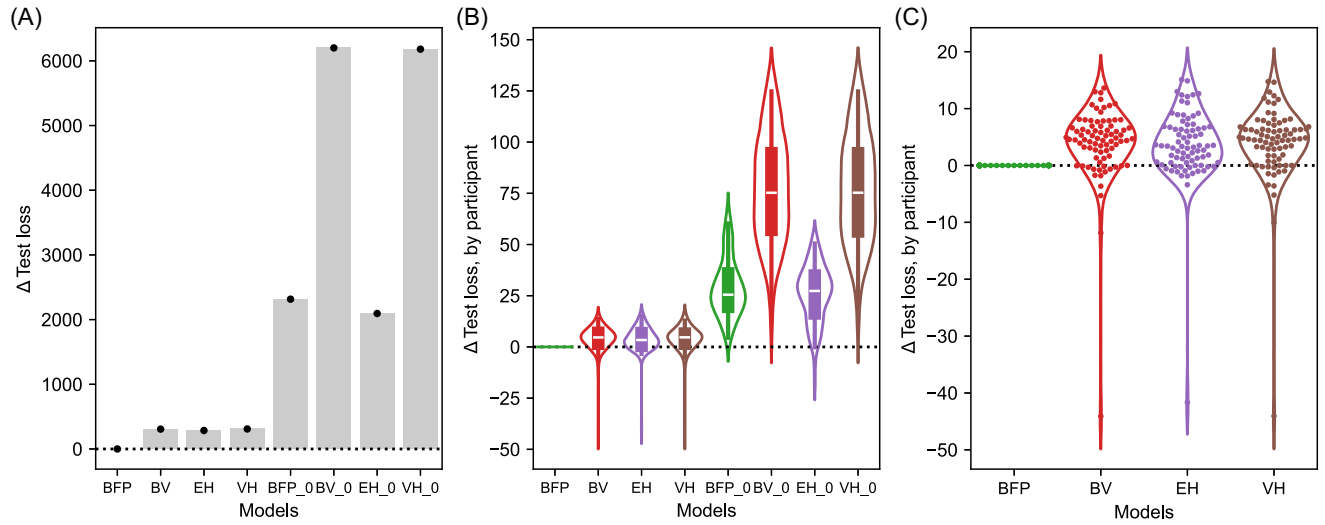
The modeling results of Experiment 3 differ from Experiments 1 and 2 in meaningful ways. In Experiments 1 and 2, even though the BFP model with logistic offset  $b$  was the winning model, its counterpart without  $b$  also outperformed the alternative models. However, in Experiment 3, the logistic offset parameter became more crucial in capturing the behavioral data. As revealed through the model visualizations (Figures 6 and 7) and model comparison (Figure 8), models without  $b$  predicted much more frequent switching at earlier times, which was reasonable because the performance trajectories had not separated (thus leading to chance levels of switching without decision bias). Such chance-level switching is also a potential reason why model recovery was poorer for the models without  $b$ —they encouraged switching early on in a run, leaving very few data points to fit models over (see Figure A2).

In contrast to the models without  $b$ , participants tended to stay when evidence was ambiguous before trajectories separated, leading to improved fits of models with  $b$ —which superficially reflects a decision bias toward staying. This offset may reflect a *stay-bias* arising from factors such as perseverative tendencies, uncertainty aversion (as switching introduces additional uncertainty regarding final asset performance), or past-looking sunk-cost considerations. Even if sunk costs involve a retrospective component, any such past-

**Figure 7**  
Experiment 3 Human Behavior and Modeling Results Continued



**Note.** The models visualized have the logistic offset  $b$  fixed to 0 (no decision bias). (A–C) Data (black) and model-predicted (color) mean switch time (A), mean starting asset performance at switch (B), and probability of continuing with the starting asset till the end of the run (C), for the “separate-late” and “separate-early” task conditions. (D) The starting asset’s true final performance (if the participant had continued till the end of the run), for runs where participants continued till the end versus switched midway. (E–G) Data (black) and model-predicted (color) probability of staying at each time point given that they have not switched so far in the run, visualized for both task conditions together (E), the “separate-late” condition (F), and “separate-early” condition (G). VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

**Figure 8***Experiment 3 Test Losses for Cross-Validation*

**Note.** (A) Cross-validation fold-averaged test loss differences across models. The visualized losses are summed across the individual participant fits. The error bars denote the standard error of the mean of test losses over participants. The “0” suffix denotes models fitted without the logistic offset parameter  $b$ . (B) Distribution of individual fold-averaged test losses across all candidate models. Error bars within each violin plot denote the median and interquartile ranges. (C) Zoom-in plot for the four models with logistic offset fitted. Points represent individual participants. VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

looking influence appears to be outweighed by future-looking considerations, as evidenced by the superior fit of the BFP model relative to the EH model (which focuses on asset performance achieved so far and should thus capture sunk costs). Alternatively, the decision offset may reflect future-looking reasoning—participants might be saving their only opportunity to switch to a later time point, when plausible trajectories would be separated more. Such *opportunity-saving* would require inferring latent growth structure and anticipating that plausible trajectories that have not yet diverged will become discernible in the future. Both of these are plausible cognitive interpretations of the logistic offset. Because we did not allow participants to switch more than once in each run, we were unable to distinguish between the two.

To fully distinguish the stay-bias and opportunity-saving hypotheses, we would need a task that allows participants to switch an unlimited number of times per run. On the modeling side, this complicates the picture because the utility of switching must now incorporate the recursive utilities at later times associated with switching and not switching then. Future modeling studies could address this shortcoming and lead to a more realistic account of human asset switching.

### Experiment 4

In Experiment 4, we examine whether Bayesian forecasting generalizes to a broader range of performance dynamics. Specifically, we extended the framework from monotonic logistic growth (in Experiments 1–3) to boom-and-bust trajectories, a classical functional form characterized by initial growth, a performance peak, and eventual decline (Diebold & Rudebusch, 1994). Such trajectories are common across many real-world settings including skill acquisition under fatigue or forgetting (e.g., preparing before a competition or performance; Branscheidt et al., 2019), campaigns or outreach efforts that

peak in engagement before attention wanes (e.g., political mobilization before election day; Gerber et al., 2011; Panagopoulos, 2011), and biological growth processes (e.g., population changes or muscle growth under overtraining; Kenttä & Hassmén, 1998; Strayer et al., 2017). In all these domains, performance at a fixed evaluation point—such as the competition/interview date, election day, or suitable season in the year—serves as the ultimate determinant of success.

When the growth phase of typical boom-and-bust cycles is shorter than the anticipated time horizon, the central decision problem shifts to when to initiate development, such that the asset’s performance peaks precisely at the evaluation time. To operationalize this idea within the same decision framework as Experiments 1–3, we reformulated the problem as asset switching: Given a starting asset that has already entered its boom-and-bust cycle, participants must decide at every timestep whether to continue developing it or switch to a new asset whose growth cycle has yet to begin. This reframing preserves the experimental consistency of the continue-switch paradigm while probing the same underlying computational challenge—forecasting future trajectories and maximizing final performance at some designated time horizon. Experiment 4 explores whether the BFP model can account for human asset switching behavior when asset performance follows transient, nonmonotonic dynamics—a domain in which linear extrapolation into the future would lead them astray.

## Materials and Method

### Participants

A total of 80 participants (32 women, 48 men) were recruited on MTurk. We selected the sample size based on a pilot experiment ( $N = 30$ ) to detect the effects of interest with 80% power. Participants



received a \$5 base pay and a bonus up to \$2 proportional to the mean final performance achieved across test runs. Participants were allowed to participate in the main experiment only after passing a comprehension check following the instructions. We did not exclude any participant or observation. Participants gave informed consent, and the experiment was approved by the Harvard Institutional Review Board and preregistered at <https://aspredicted.org/c8nz6a.pdf>.

### Stimuli

The same cover story of defense systems as assets was used. The trajectory distributions differed across the two conditions, being piecewise instead of logistic as in previous experiments:

$$\rho(t; \theta_i) = \begin{cases} \theta_i^0 + \frac{1}{\theta_i^1} \cdot t, & \text{if } t \leq \theta_i^1 \\ \frac{\theta_i^1 \cdot \theta_i^2 - 10}{\theta_i^1 - 10} + \frac{1 - \theta_i^2}{\theta_i^1 - 10} \cdot t, & \text{otherwise} \end{cases}, \quad (12)$$

where  $\theta_i^0 = 0$ ,  $\theta_i^1 = (\alpha + \beta)$ ,  $\theta_i^2 = (0.5 + 0.2\alpha)$ , and  $\alpha \sim N(0, 0.2^2)$  for both conditions. For the “peak-early” condition,  $\beta = 2$ ; for the “peak-late” condition,  $\beta = 8$ . In other words, the trajectory distributions  $p(\theta)$  used for the two conditions would start from zero performance, increase to their peak performance of one unit at around  $t = 2$  or  $8$  respectively, and decrease afterward. Stochasticity was introduced by the single trajectory parameter  $\alpha$  (Figure 9A). Our experiment directly visualized the above  $\rho(t; \theta_i)$ , leading to the simplification  $r_i^{(t)} = \rho(t; \theta_i)$  in the true generative process. Akin to Experiment 3, the trajectories used in training and test time were randomly sampled for each participant in each run. The starting asset is always displayed above the other asset that could be switched into.

Similar to Experiment 3, an asset’s performance was conveyed through the total length of a corresponding horizontal green bar (Figure 9B). To represent asset performance decrements in Experiment 4, we truncated the green bars from their right end by the appropriate length, such that their remaining total length accurately conveyed current asset performance (Figure 9C). To prevent participant confusion, we refrained from displaying black vertical lines segmenting growth at each timestep, for both performance increments and decrements.

The response interface was identical to Experiment 3. Participants pressed keyboard keys—the “right” arrow to continue developing the current asset and the “down” arrow to switch to a new asset. They could switch at most once per run.

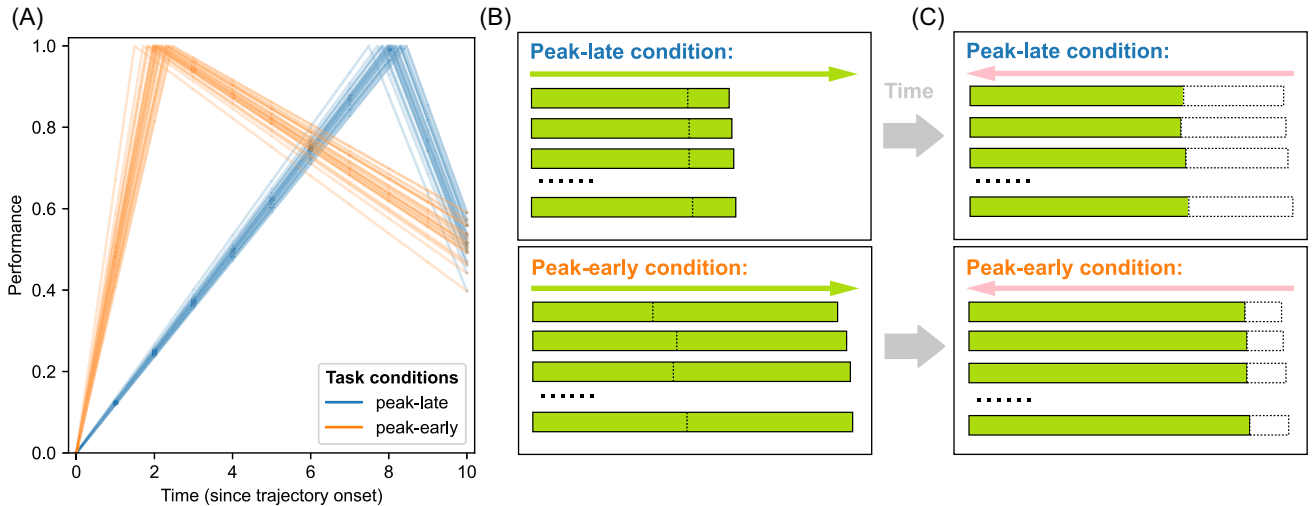
### Procedure

The procedure was identical to Experiment 3, with comprehension questions as well as condition-specific training rounds and test rounds. In the instructions, participants were only told that asset performance typically first grows and then declines, and had to infer condition-specific trajectory distribution properties from the training block examples. The only differences from Experiment 3 are the condition-specific trajectory distributions and that each test block contained 30 runs.

### Statistical Analyses

We conducted two-sided paired  $t$  tests. In the main text, we report the  $t$ -statistics, their corresponding  $p$  values, the Cohen’s  $d$  effect sizes, and 95% CIs of the effect sizes. The prediction on condition-

**Figure 9**  
Experiment 4 Setup, Performance Trajectories, and Display



**Note.** (A) Example performance trajectories sampled from the “peak-late” (blue) and “peak-early” (orange) conditions. (B) The display of performance trajectories as they increase in performance, during earlier timesteps in training. The gradually lengthening green bars illustrate how the 50 training trajectories were displayed—rows correspond to different assets sampled from the same trajectory distribution, and they are ordered from lowest to highest peak performance. Subjects could press keyboard keys to toggle back and forth in time to see how all these training trajectories develop, until they felt ready to begin the task. (C) The display of performance trajectories as they decrease in performance, during later timesteps in training. The total length of the green bars is truncated from the right. See the online article for the color version of this figure.

specific mean switch times was assessed based on all participants who switched at least once in both conditions ( $N = 76$ ).

### Model Fitting and Comparison

The model-fitting pipeline was identical to Experiment 3, using the PyBADs optimizer with 10 random parameter initializations per fit. The fivefold cross-validation pipeline for model comparison was also identical.

### Results

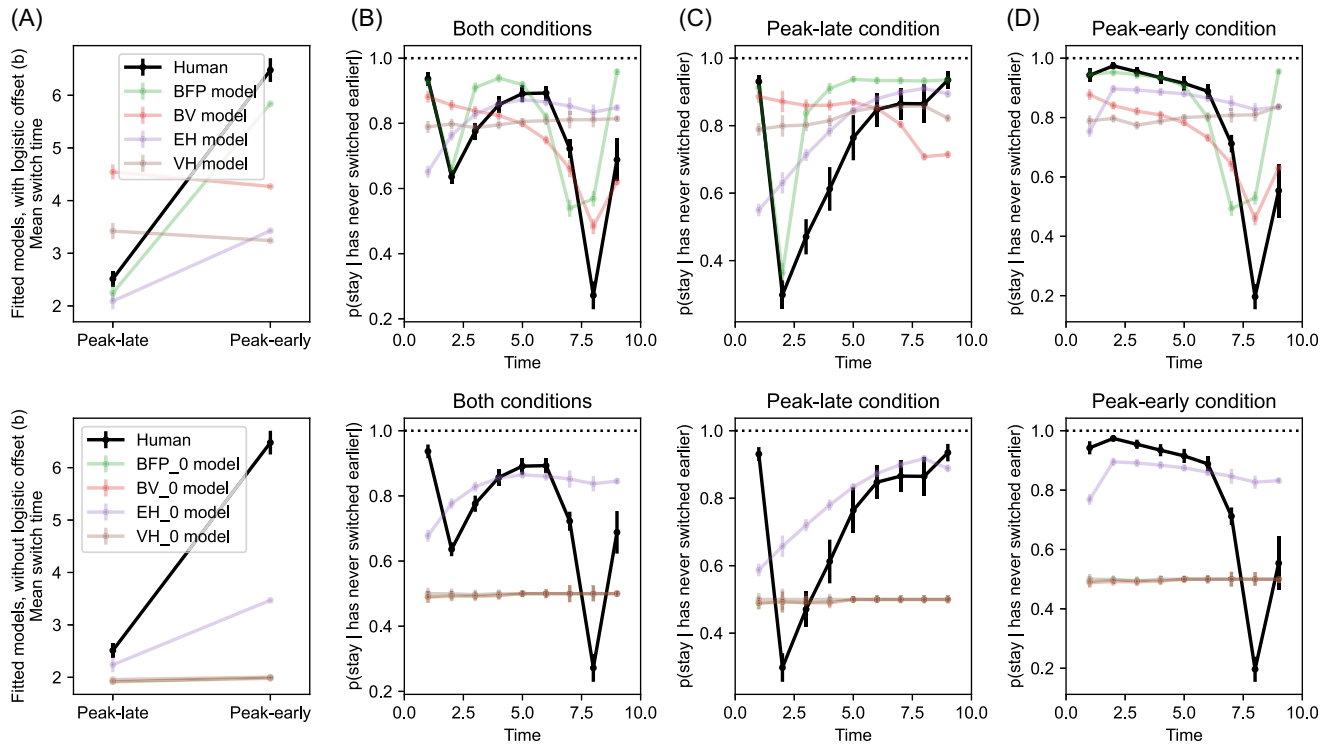
Across both task conditions, the mean number of analyzable stay-switch responses provided by participants was  $295 \pm 9.18$ , excluding the  $t = 0$  and  $t = 10$  times of every run and all the time points within a run after switching because participants could not choose during these times. Among the 80 participants, 76 switched at least once in both conditions. The number of data points per participant and their variability enabled the fitting of participant-specific models and their analysis.

Consistent with the BFP model predictions, participants switched significantly later in the peak-early condition than in the peak-late condition,  $t(75) = -12.9$ ,  $p < 10^{-20}$ ,  $d = 1.483$ , 95% CI  $[-4.58, -3.35]$  (Figure 10A). The alternative models could not replicate this pattern: The fitted BV and VH models predicted the opposite effect,

whereas the EH model failed to reproduce the magnitude of the timing difference. Examination of switch probabilities at each timestep further revealed that only the BFP model with a logistic offset captured the qualitative bimodal switching pattern observed in human data—peaking around  $t = 2$  in the peak-late condition and around  $t = 8$  in the peak-early condition (Figure 10B–10D, Row 1). This pattern reflected the intuitive human tendency to switch at a timestep just enough for the new asset to reach its performance peak. However, quantitative disparities do exist. For the peak-late condition, participants exhibited stronger tendencies to switch after  $t = 2$  if they had failed to switch before then. For the peak-early condition, the BFP model predicted stronger switching tendencies at  $t = 7$  than  $t = 8$  given the precise trajectory distribution, while participants preferred to switch later at  $t = 8$  instead.

The logistic offset parameter  $b$  is again crucial in capturing human behavior—without it, the BFP\_0, BV\_0, and VH\_0 models favor switching predominately. After all, switching after  $t = 0$  would consistently lead to a shorter decline phase from peak performance compared to staying with the current asset until the end. To account for the fact that participants mostly continued and only switched selectively at different timesteps across conditions, these models converged to near-random predictions across all timesteps to minimize training loss, which still deviated prominently from human behavior (Figure 10B–10D, Row 2).

**Figure 10**  
*Experiment 4 Human Behavior and Modeling Results*



**Note.** Error bars denote the standard error of the mean across participants. *Row 1:* Fitted models with the logistic offset  $b$  as a free parameter. *Row 2:* Fitted models without the logistic offset  $b$ . (A) Mean switch time of human data (black) and model predictions (color) for each task condition. (B–D) Human (black) and model-predicted (color) probability of staying at each time point given that they have not switched so far in the run, visualized for both task conditions together (B), the “peak-late” condition (C), and “peak-early” condition (D). VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

Model comparison via fivefold cross-validation similarly favored the BFP model with logistic offset (Figure 11A). Again, this conclusion is not the result of a few outliers: Two-sided paired  $t$  tests between the BFP model with logistic offset  $b$  and the other models with or without  $b$  are all significant at the  $\alpha = .0001$  level (Figure 11B and 11C). Model recovery supports the validity of the above procedure. Excluding models that produced nearly identical predictions across timesteps (and could thus be mimicked by any model with sufficiently large decision noise and a suitable offset), cross-validation reliably distinguished among the BFP, BV, EH, and VH models (Figure B1).

## Discussion

Experiment 4 follows Experiment 3 in providing a naturalistic task setup in which participants were allowed to switch at any time. Again, the BFP model—forecasting final asset performances associated with continuing till the end versus switching—was the only model that could capture human switching patterns across the two task conditions. The selective switching of participants at  $t = 2$  for the peak-late condition and  $t = 8$  for the peak-early condition reveals context-dependent timing that, again, could not be captured by the fixed-time foraging heuristic. Similarly, the fixed-number foraging heuristic—prescribing switching once the current asset’s performance reaches a certain threshold—could not account for human switching at  $t = 8$  in the peak-early condition, long after the starting asset’s performance had begun to decline. In this task condition, from  $t = 0$  to 3, the starting asset’s performance would have already traversed its entire possible range during its growth phase (from zero to one unit), such that a fixed-number agent would have switched before the peak no matter what their threshold might be.

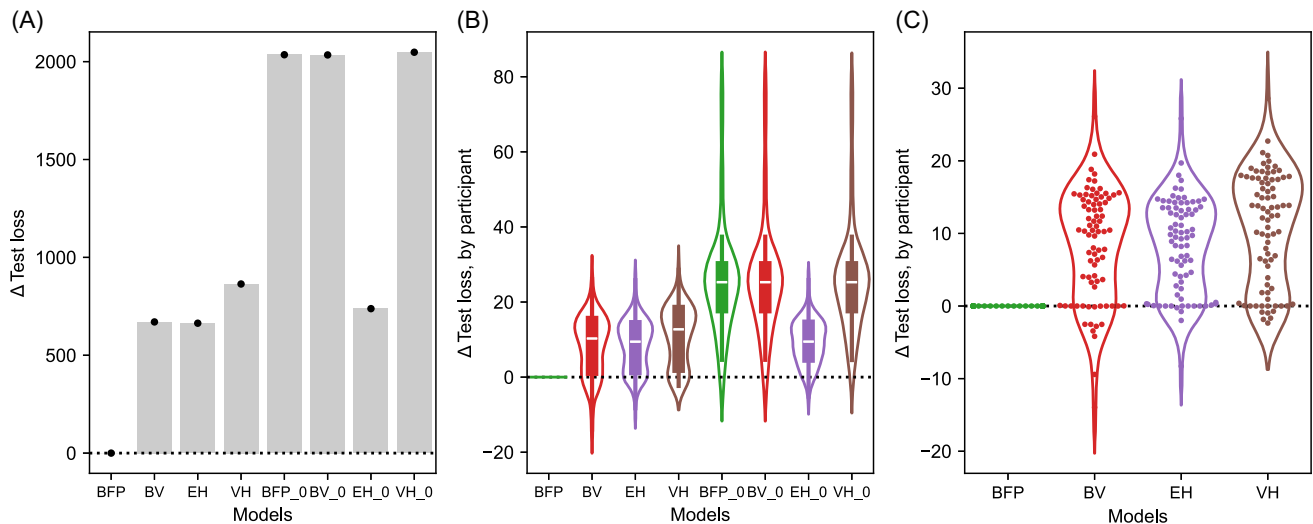
The fact that participants switched at around the optimal time prescribed by Bayesian forecasting highlights their capability to make forecasts beyond linearly extrapolating past trends. Instead, to perform well in Experiment 4, humans would need to at least store the sufficient statistic on how much time is required for an asset to reach peak performance, if not more information on the rate of growth and decline. However, the memory demands of doing so—along with tracking the current time and asset performance—likely prevented a precise implementation of forecasting and comparison as specified in the normative BFP model. This may underlie the quantitative mismatches observed between humans and BFP and raises the question of plausible process-level approximations participants may be using.

The necessity of the logistic offset parameter  $b$  in capturing human behavior further reinforces that it functions as more than a simple, context-independent bias to remain with the current asset. As in Experiment 3—where the offset delayed premature switching until trajectories within the same distribution began to diverge, thus preserving the opportunity to switch when it would have the greatest impact—in Experiment 4, it similarly delays switching to the timestep where it could lead to the greatest improvement in final performance ( $t = 2$  for the peak-late condition and  $t = 8$  for peak-early). This pattern again suggests that humans may engage in recursive reasoning about the value of switching now versus conserving the option to switch later, highlighting an important direction for future modeling work.

## General Discussion

In this article, we present a systematic study of asset selection and switching problems. This problem domain is characterized by assets that develop with investment (which tend to improve over time but

**Figure 11**  
Experiment 4 Test Losses for Cross-Validation



**Note.** (A) Cross-validation fold-averaged test loss differences across models. The visualized losses are summed across the individual participant fits. The error bars denote the standard error of the mean of test losses over participants. The “0” suffix denotes models fitted without the logistic offset parameter  $b$ . (B) Distribution of individual fold-averaged test losses across all candidate models. Error bars within each violin plot denote the median and interquartile ranges. (C) Zoom-in plot for the four models with logistic offset fitted. Points represent individual participants. VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

not necessarily), assets that can only be developed one at a time, and the time horizons at which only the current asset's performance is evaluated. This problem resembles a broad class of important real-life scenarios, from deciding which project to take on and which trainee to recruit to whether and when to give up a project, quit a job, or leave a relationship. Navigating these problems requires probabilistic forecasting of how assets will develop while taking into consideration the relevant time horizons. These features make the myopic solutions in the fields of foraging and study time allocation less likely to apply.

Here, we develop a normative Bayesian framework for asset selection and asset switching problems. The framework assumes that people extrapolate the performance trajectories of assets into the future and consider them in conjunction with time horizons, with its plausibility supported by previous work in function learning and evidence accumulation. Across three experiments, we manipulated whether there was prior investment, the trajectory shapes and time horizons, and the degree of freedom in making the decisions. We demonstrated that the Bayesian modeling framework captures participants' behavior better than alternative models inspired by foraging and study time allocation, as well as a Bayesian model based on future improvement velocity. These results suggest that people use probabilistic forecasting to guide decisions with long-term implications, as opposed to myopic heuristics alone.

The present study advances Bayesian function learning in humans to more complex domains and situates it within the broader context of mental models in human model-based RL. Previously, researchers have mostly studied human function learning either as an isolated cognitive capability (C. G. Lucas et al., 2015) or used in conjunction with decision making to estimate the utility of unexplored options (Schulz et al., 2018; Wu et al., 2018). In contrast, there have been limited studies that probe function learning in temporally extended decisions. Our work bridges this gap by demonstrating that Bayesian function learning naturally supports time-dependent decision making and confirming its relevance to human behavior. Its core usage—function approximation for prediction—may represent a critical yet underexplored component of model-based learning: the ability to internally simulate how actions or merely time could transform environmental dynamics and rewards. By emphasizing this predictive function-learning process, we conceptually connect human function learning with model-based RL and highlight how people use mental models of change over time to guide decision making.

More generally, the success of the Bayesian forecasting model reveals a key theoretical insight: Human decision makers are future-looking, beyond merely expecting past trends to continue (Harvey & Reimers, 2013; Lawrence et al., 2006; Sanders & Ritzman, 1992, 1995). In line with the Lucas critique in economics (R. E. Lucas, 1976), humans appear to account for how their own choices would reshape the environmental dynamics that determine future rewards—where assets such as employees, projects, jobs, skills, or even interpersonal relationships improve or fluctuate with additional practice or commitment (Xiang & Gershman, 2025). These anticipatory forecasts are in turn fed back to their present choices, which embody long-term consequences as characterized above. This future-looking tendency stands in contrast to more traditional frameworks such as evidence accumulation, patch foraging, and model-free RL. These classical frameworks instead assume a relatively static environment and emphasize directly integrating past outcomes, caching state–action–

reward associations, or evaluating only the current state. Holistically, the empirical results suggest that Bayesian forecasting provides a principled computational account of future-looking cognition—one that captures how humans integrate prediction and action under nonstationary conditions, where their own choices have lasting consequences.

## Limitations and Future Directions

Our work provides a first step to understanding dynamic forecasting in its basic form. Real-world asset decisions, however, often involve features that are not present in our tasks, including but not limited to multiple evaluation milestones, heterogeneous deadlines, monetary or time costs for switching, parallel investment across options, or cognitive constructs such as satisficing behavior, subjective utility transformations, temporal discounting, and risk preferences. While the Bayesian forecasting component already captures a substantial proportion of behavioral variance, it remains an exciting future direction to examine how forecasting interacts with the above realistic considerations.

The framework we propose can be extended to predict the trajectory of switching among assets if we allow switching multiple times, and potentially switching back to a previously invested asset. The possibility of switching back may be especially helpful when there is greater uncertainty with how assets develop and when assets develop unexpectedly (e.g., a seemingly easy project becomes much harder than expected). However, it requires calculating utilities recursively, by anticipating not only the utilities of staying or switching now but also the possibility of switching again later, which can quickly become computationally intractable. One solution is to combine RL and Bayesian function learning mechanisms to provide a more comprehensive account of how people make these decisions (Schulz et al., 2018; Wu et al., 2018).

We made a number of simplifications in our experiments. First, we always included a training phase that allowed participants to develop priors over the performance trajectories. However, in real life, the experiences become the new priors—humans must learn priors as they observe assets, and simultaneously make selection and switching decisions to maximize final performance. Hence, when an inexperienced agent with weak priors unluckily starts with a poorly performing asset, they may erroneously infer that the entire trajectory distribution is poor and be discouraged from switching, an explanation that rationalizes “sunk costs” usually thought to be irrational (McAfee et al., 2010). Second, we assumed shared time horizons across assets in our experiments, but this is not a core assumption of the framework. The framework can be straightforwardly extended to deal with situations where the time horizon may change following a switch, whether due to general environmental reasons or asset-specific characteristics. The BFP model can flexibly use asset-specific time horizons to forecast their final performance. However, it still leaves open the question of how people infer the relevant time horizons in the real world (Griffiths & Tenenbaum, 2006, 2013). Inaccurate priors over time horizons might explain human suboptimalities in asset selection and switching (McGuire & Kable, 2012). Third, while we focus on the chosen asset's performance at a specific time horizon, some asset selection/switching problems may instead involve maximizing cumulative rewards over time (e.g., choosing jobs to optimize total savings by retirement). To better understand human decision making in these scenarios, it is important to examine the environmental conditions



under which probabilistic forecasting provides an advantage over myopic heuristics, develop corresponding decision models, and assess whether such forecasting is integrated into human strategies.

Our study also raises several questions for future investigations. While making asset selection and switching decisions, people are simultaneously investing effort into their current asset, which may take up cognitive resources. How do we navigate both the low-level action and the metacognitive decision? For example, if the current job is very exhausting or requires extensive attention engagement, it would be hard to notice or consider other jobs while maintaining the current job (Holton et al., 2024). In these situations, would people ignore available alternatives along with the possibility of switching or resort to myopic heuristics more? Evidence from the study time allocation literature suggests that high-level metacognition is still at play when humans solve low-level tasks (Payne et al., 2007), and they have a basic understanding of typical learning trajectory shapes needed for Bayesian priors (Zhang et al., 2025). This suggests that people may still be able to consider asset selection and switching while engaged in lower level tasks, but exactly how these low-level actions affect selection and switching decisions is an open question for future work.

Given resource constraints, a natural next question concerns whether people rely on heuristics to approximate Bayesian forecasting on the process level. Previous research has documented a tendency to use linear extrapolation to forecast ongoing trends (Andreassen, 1990; DeLosh et al., 1997; Harvey et al., 1994). However, such a strategy alone cannot account for human behavior across all four experiments. As discussed in Experiments 2 and 4, the empirical patterns would differ substantially if participants had used purely linear extrapolation that disregards the characteristic shapes of logistic growth or regressive trajectories. An alternative possibility is that people draw from an inventory of heuristics that are selectively deployed depending on the typical distribution of asset-development trajectories in a given task, analogous to the adaptive toolbox proposed in ecological rationality (Gigerenzer & Selten, 2002). Indeed, rule-based or conditional (“if-then”) architectures could plausibly capture participant behavior (Goodman et al., 2008; Shea, 2024). For example, in Experiment 3, a participant might follow a rule such as “switch if and only if the condition is separate-early and the asymptotic performance of the starting asset falls below a threshold”; in Experiment 4, a simpler heuristic such as “switch at timestep ( $T - t_{\text{peak}}$ )” might be used. A key open question is how humans decide which heuristic to deploy in a given task environment (combination of trajectory distributions and time horizons) to achieve near-optimal or satisficing performance. Although the Bayesian forecasting framework is normative rather than descriptive, it provides a valuable benchmark for understanding how people might acquire, evaluate, and manage the selective use of heuristics across different contexts.

A third question concerns asset selection and switching when more than two alternatives are available. As discussed in the introduction, this includes scenarios where the agent must choose from three or more assets or has the flexibility to switch to a specific asset among many options. Expanding the framework to accommodate these cases may require not only new evidence accumulation models (Brown & Heathcote, 2008; Tajima et al., 2019) but also a consideration of the cognitive load associated with forecasting and comparing multiple assets, as discussed above. Previous research has shown that the number of available choices influences switching

behavior in the context of static-valued goods (Ratner & Herbst, 2005) and that humans may spontaneously simplify tasks by considering only a subset of available options under cognitive load (Liu & Gershman, 2025). Additionally, related work in foraging suggests that the flexibility to switch to particular alternatives may alter human decision strategies (Hall-McMaster & Luyckx, 2019). Given such neighboring evidence, future research should examine how the presence of multiple alternatives influences human asset selection and switching.

Last, the development of an asset is also closely related to how we invest in it. People usually have some control over the growth of an asset; for example, putting more effort into a research project can lead to more progress, and treating one’s partner with love and respect can foster a better relationship. Indeed, past work has shown that, when training collaborators, we need to decide not just whom to train but also how much to train them (Oudeyer & Kaplan, 2007; Tian et al., 2025; Xiang et al., 2024). The right amount of investment also depends on the performance of the asset; while we showed participants the asset’s performance through the length of a bar, this information often has to be inferred from observations (Xiang et al., 2023). The current work is limited in that participants did not have control over how the trajectory develops. Future experiments should examine how people make these asset selection and switching decisions when they have agency in investing in the assets and, conversely, how the option to switch out affects their investment in the current asset.

## References

- Acerbi, L., & Ma, W. J. (2017). Practical bayesian optimization for model fitting with bayesian adaptive direct search. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (pp. 1834–1844). Curran Associates. <https://doi.org/10.48550/arXiv.1705.04405>
- Andreassen, P. B. (1990). Judgmental extrapolation and market overreaction: On the use and disuse of news. *Journal of Behavioral Decision Making*, 3(3), 153–174. <https://doi.org/10.1002/bdm.3960030302>
- Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, 13(3), 219–224. <https://doi.org/10.1111/1467-9280.00441>
- Baumann, C., Singmann, H., Gershman, S. J., & von Helversen, B. (2020). A linear threshold model for optimal stopping behavior. *Proceedings of the National Academy of Sciences*, 117(23), 12750–12755. <https://doi.org/10.1073/pnas.2002312117>
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), Article 700. <https://doi.org/10.1037/0033-295X.113.4.700>
- Branscheidt, M., Kassavetis, P., Anaya, M., Rogers, D., Huang, H. D., Lindquist, M. A., & Celnik, P. (2019). Fatigue induces long-lasting detrimental changes in motor-skill learning. *eLife*, 8, Article e40578. <https://doi.org/10.7554/eLife.40578>
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3), 153–178. <https://doi.org/10.1016/j.cogpsych.2007.12.002>
- Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9(2), 129–136. [https://doi.org/10.1016/0040-5809\(76\)90040-X](https://doi.org/10.1016/0040-5809(76)90040-X)
- Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4), 404–413. <https://doi.org/10.1093/biomet/26.4.404>

- Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, 15, 837–853. <https://doi.org/10.3758/s13415-015-0350-y>
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4), 429–453. <https://doi.org/10.3758/CABN.8.4.429>
- De Bondt, W. P. (1993). Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting*, 9(3), 355–371. [https://doi.org/10.1016/0169-2070\(93\)90030-Q](https://doi.org/10.1016/0169-2070(93)90030-Q)
- De Rooij, M., & Weeda, W. (2020). Cross-validation: A method every psychologist should know. *Advances in Methods and Practices in Psychological Science*, 3(2), 248–263. <https://doi.org/10.1177/2515245919898466>
- DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 968–986. <https://doi.org/10.1037/0278-7393.23.4.968>
- Diebold, F. X., & Rudebusch, G. D. (1994). *Measuring business cycles: A modern perspective*. National Bureau of Economic Research. <https://doi.org/10.3386/w4643>
- Fontanesi, L., Gluth, S., Spector, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, 26(4), 1099–1121. <https://doi.org/10.3758/s13423-018-1554-2>
- Gerber, A. S., Gimpel, J. G., Green, D. P., & Shaw, D. R. (2011). How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. *American Political Science Review*, 105(1), 135–150. <https://doi.org/10.1017/S000305541000047X>
- Gershman, S. J., Malmaud, J., & Tenenbaum, J. B. (2017). Structured representations of utility in combinatorial domains. *Decision*, 4(2), 67–86. <https://doi.org/10.1037/dec0000053>
- Gershman, S. J., & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science*, 7(3), 391–415. <https://doi.org/10.1111/tops.12138>
- Gershman, S., Cohen, J., & Niv, Y. (2010). Learning to selectively attend. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 32, 1270–1275. <https://escholarship.org/uc/item/7bn072ds#main>
- Ghosh, B., & Crowley, M. (2019). *The theory behind overfitting, cross validation, regularization, bagging, and boosting: Tutorial*. arXiv preprint.
- Gibb, J. A. (1962). L. Tinbergen's hypothesis of the role of specific search images. *Ibis*, 104(1), 106–111. <https://doi.org/10.1111/j.1474-919X.1962.tb08633.x>
- Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. MIT Press. <https://doi.org/10.7551/mitpress/1654.001.0001>
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 351–360. [https://doi.org/10.1016/S0022-5371\(66\)80044-0](https://doi.org/10.1016/S0022-5371(66)80044-0)
- Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science*, 32(1), 108–154. <https://doi.org/10.1080/03640210701802071>
- Goodwin, P., & Wright, G. (1993). Improving judgmental time series forecasting: A review of the guidance provided by research. *International Journal of Forecasting*, 9(2), 147–161. [https://doi.org/10.1016/0169-2070\(93\)90001-4](https://doi.org/10.1016/0169-2070(93)90001-4)
- Green, R. F. (1984). Stopping rules for optimal foragers. *American Naturalist*, 123(1), 30–43. <https://doi.org/10.1086/284184>
- Griffiths, T. L., Chater, N., & Tenenbaum, J. B. (2024). *Bayesian models of cognition: Reverse engineering the mind*. MIT Press.
- Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science*, 17(9), 767–773. <https://doi.org/10.1111/j.1467-9280.2006.01780.x>
- Griffiths, T. L., & Tenenbaum, J. B. (2013). Predicting the future as bayesian inference: People combine prior knowledge with observations when estimating duration and extent. *Decision*, 1, 56–82. <https://doi.org/10.1037/2325-9965.1.S.56>
- Guan, M., Lee, M., & Silva, A. (2014). Threshold models of human decision making on optimal stopping problems in different environments. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36, 553–558. <https://escholarship.org/uc/item/6k95q5hb>
- Hall-McMaster, S., Dayan, P., & Schuck, N. W. (2021). Control over patch encounters changes foraging behavior. *iScience*, 24(9). <https://doi.org/10.1016/j.isci.2021.103005>
- Hall-McMaster, S., & Luyckx, F. (2019). Revisiting foraging approaches in neuroscience. *Cognitive, Affective, & Behavioral Neuroscience*, 19, 225–230. <https://doi.org/10.3758/s13415-018-00682-z>
- Harvey, N., Bolger, F., & McClelland, A. (1994). On the nature of expectations. *British Journal of Psychology*, 85(2), 203–229. <https://doi.org/10.1111/j.2044-8295.1994.tb02519.x>
- Harvey, N., & Reimers, S. (2013). Trend damping: Under-adjustment, experimental artifact, or adaptation to features of the natural environment? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(2), 589–607. <https://doi.org/10.1037/a0029179>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Hohle, S. M., & Teigen, K. H. (2015). Forecasting forecasts: The trend effect. *Judgment and Decision Making*, 10(5), 416–428. <https://doi.org/10.1017/S1930297500005568>
- Holton, E., Grohn, J., Ward, H., Manohar, S. G., O'reilly, J. X., & Kolling, N. (2024). Goal commitment is supported by vmPFC through selective attention. *Nature Human Behaviour*, 8(7), 1351–1365. <https://doi.org/10.1038/s41562-024-01844-5>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- Johnson, S. G., & Tuckett, D. (2022). Narrative expectations in financial forecasting. *Journal of Behavioral Decision Making*, 35(1), Article e2245. <https://doi.org/10.1002/bdm.2245>
- Jones, G. V. (1979). A generalized polynomial model for perception of exponential series. *Perception & Psychophysics*, 25(3), 232–234. <https://doi.org/10.3758/BF03202992>
- Kentä, G., & Hassmén, P. (1998). Overtraining and recovery: A conceptual model. *Sports Medicine*, 26(1), 1–16. <https://doi.org/10.2165/00007256-199826010-00001>
- Kilpatrick, Z. P., Davidson, J. D., & El Hady, A. (2021). Uncertainty drives deviations in normative foraging decision strategies. *Journal of the Royal Society Interface*, 18(180), Article 20210337. <https://doi.org/10.1098/rsif.2021.0337>
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Vol. 14, pp. 1137–1145). <https://dl.acm.org/doi/10.5555/1643031.1643047>
- Krebs, J. R. (1973). Behavioral aspects of predation. In P. P. G. Bateson & P. H. Klopfer (Eds.) *Perspectives in ethology* (pp. 73–111). Springer. [https://doi.org/10.1007/978-1-4615-7569-6\\_3](https://doi.org/10.1007/978-1-4615-7569-6_3)
- Lawrence, M., Goodwin, P., O'Connor, M., & Önköl, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22(3), 493–518. <https://doi.org/10.1016/j.ijforeca.2006.03.007>
- Lee, M. D. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, 30(3), 1–26. [https://doi.org/10.1207/s15516709cog0000\\_69](https://doi.org/10.1207/s15516709cog0000_69)
- Li, M. Y., Callaway, F., Thompson, W. D., Adams, R. P., & Griffiths, T. L. (2023). Learning to learn functions. *Cognitive Science*, 47(4), Article e13262. <https://doi.org/10.1111/cogs.13262>

- Li, M. Y., Shi, C., Wu, Z., & Fryzlewicz, P. (2025). Testing stationarity and change point detection in reinforcement learning. *The Annals of Statistics*, 53(3), 1230–1256. <https://doi.org/10.1214/25-AOS2501>
- Lin, H.-Y., & von Helversen, B. (2023). Never gonna give you up even when it is suboptimal. *Cognitive Science*, 47(7), Article e13323. <https://doi.org/10.1111/cogs.13323>
- Liu, S. (2026). *Forecasting guides dynamic decisions* [Data set]. GitHub.
- Liu, S., & Gershman, S. J. (2025). Action subsampling supports policy compression in large action spaces. *PLOS Computational Biology*, 21(9), Article e1013444. <https://doi.org/10.1371/journal.pcbi.1013444>
- Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic Bulletin & Review*, 22(5), 1193–1215. <https://doi.org/10.3758/s13423-015-0808-5>
- Lucas, R. E., Jr. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* (Vol. 1, pp. 19–46). North-Holland. [https://doi.org/10.1016/S0167-2231\(76\)80003-6](https://doi.org/10.1016/S0167-2231(76)80003-6)
- Ma, W. J., Kording, K. P., & Goldreich, D. (2023). *Bayesian models of perception and action: An introduction*. MIT Press.
- McAfee, R. P., Mialon, H. M., & Mialon, S. H. (2010). Do sunk costs matter? *Economic Inquiry*, 48(2), 323–336. <https://doi.org/10.1111/j.1465-7295.2008.00184.x>
- McGuire, J. T., & Kable, J. W. (2012). Decision makers calibrate behavioral persistence on the basis of time-interval experience. *Cognition*, 124(2), 216–226. <https://doi.org/10.1016/j.cognition.2012.03.008>
- McNair, J. N. (1982). Optimal giving-up times and the marginal value theorem. *American Naturalist*, 119(4), 511–529. <https://doi.org/10.1086/283929>
- Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, 52(4), 463–477. <https://doi.org/10.1016/j.jml.2004.12.001>
- Murre, J. M. (2014). S-shaped learning curves. *Psychonomic Bulletin & Review*, 21, 344–356. <https://doi.org/10.3758/s13423-013-0522-0>
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3), 139–154. <https://doi.org/10.1016/j.jmp.2008.12.005>
- Oudeyer, P.-Y., & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurobotics*, 1, Article 2007. <https://doi.org/10.3389/neuro.12.006.2007>
- Padakandla, S., Prabuchandran, K. J., & Bhatnagar, S. (2020). Reinforcement learning algorithm for non-stationary environments. *Applied Intelligence*, 50(11), 3590–3606. <https://doi.org/10.1007/s10489-020-01758-5>
- Panagopoulos, C. (2011). Timing is everything? Primacy and recency effects in voter mobilization campaigns. *Political Behavior*, 33(1), 79–93. <https://doi.org/10.1007/s11109-010-9125-x>
- Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task inter-leaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology: General*, 136(3), Article 370. <https://doi.org/10.1037/0096-3445.136.3.370>
- Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, 24, 1234–1251. <https://doi.org/10.3758/s13423-016-1199-y>
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. <https://doi.org/10.1162/neco.2008.12-06-420>
- Ratner, R. K., & Herbst, K. C. (2005). When good decisions have bad outcomes: The impact of affect on switching behavior. *Organizational Behavior and Human Decision Processes*, 96(1), 23–37. <https://doi.org/10.1016/j.obhdp.2004.09.003>
- Rosa, H. (2013). *Social acceleration: A new theory of modernity*. Columbia University Press. <https://doi.org/10.7312/rosa14834>
- Sanders, N. R., & Ritzman, L. P. (1992). The need for contextual and technical knowledge in judgmental forecasting. *Journal of Behavioral Decision Making*, 5(1), 39–52. <https://doi.org/10.1002/bdm.3960050106>
- Sanders, N. R., & Ritzman, L. P. (1995). Bringing judgment into combination forecasts. *Journal of Operations Management*, 13(4), 311–321. [https://doi.org/10.1016/0272-6963\(95\)00039-9](https://doi.org/10.1016/0272-6963(95)00039-9)
- Schulz, E., Franklin, N. T., & Gershman, S. J. (2020). Finding structure in multi-armed bandits. *Cognitive Psychology*, 119, Article 101261. <https://doi.org/10.1016/j.cogpsych.2019.101261>
- Schulz, E., Konstantinidis, E., & Speekenbrink, M. (2018). Putting bandits into context: How function learning supports decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(6), 927–943. <https://doi.org/10.1037/xlm0000463>
- Schulz, E., Tenenbaum, J. B., Duvenaud, D., Speekenbrink, M., & Gershman, S. J. (2017). Compositional inductive biases in function learning. *Cognitive Psychology*, 99, 44–79. <https://doi.org/10.1016/j.cogpsych.2017.11.002>
- Shea, N. (2024). The frame problem and the if-then problem. In N. Shea (Ed.), *Concepts at the interface* (pp. 19–46). Oxford University Press. <https://doi.org/10.1093/9780191997167.003.0006>
- Singh, G. S., & Acerbi, L. (2023). *Pybads: Fast and robust black-box optimization in python*. arXiv preprint.
- Son, L. K., & Kornell, N. (2008). Research on the allocation of study time: Key studies from 1890 to the present (and beyond). In J. Dunlosky & R. Bjork (Eds.), *A handbook of memory and metamemory* (pp. 333–351). Psychology Press. <https://doi.org/10.4324/9780203805503>
- Son, L. K., & Sethi, R. (2006). Metacognitive control and optimal learning. *Cognitive Science*, 30(4), 759–774. [https://doi.org/10.1207/s15516709cog0000\\_74](https://doi.org/10.1207/s15516709cog0000_74)
- Strayer, D. L., D'Antonio, C. M., Essl, F., Fowler, M. S., Geist, J., Hilt, S., Jaric, I., Jöhnk, K., Jones, C. G., Lambin, X., Latzka, A. W., Pergl, J., Pyšek, P., Robertson, P., von Schmalensee, M., Stefansson, R. A., Wright, J., & Jeschke, J. M. (2017). Boom-bust dynamics in biological invasions: Towards an improved application of the concept. *Ecology Letters*, 20(10), 1337–1350. <https://doi.org/10.1111/ele.12822>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- Tajima, S., Drugowitsch, J., Patel, N., & Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nature Neuroscience*, 22(9), 1503–1511. <https://doi.org/10.1038/s41593-019-0453-9>
- Taylor, H. M. (1968). Optimal stopping in a markov process. *Annals of Mathematical Statistics*, 39(4), 1333–1344. <https://doi.org/10.1214/aoms/1177698259>
- Tekin, E. (2022). Can learners allocate their study time effectively? It is complicated. *Educational Psychology Review*, 34, 717–748. <https://doi.org/10.1007/s10648-021-09645-2>
- Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, 12(1), Article 5972. <https://doi.org/10.1038/s41467-021-26196-w>
- Tian, F., Gershman, S. J., & Xiang, Y. (2025). *Training collaborators for effective division of labor*. PsyArXiv. [https://doi.org/10.31234/osf.io/dxy7u\\_v1](https://doi.org/10.31234/osf.io/dxy7u_v1)
- Topel, R. H., & Ward, M. P. (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics*, 107(2), 439–479. <https://doi.org/10.2307/2118478>
- van Ravenzwaaij, D., Brown, S. D., Marley, A., & Heathcote, A. (2020). Accumulating advantages: A new conceptualization of rapid multiple choice. *Psychological Review*, 127(2), 186–215. <https://doi.org/10.1037/rev0000166>
- Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, 2(12), 915–924. <https://doi.org/10.1038/s41562-018-0467-4>

- Xiang, Y., & Gershman, S. (2025). Modeling intrinsic motivation as reflective planning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 47, 201–208. <https://escholarship.org/uc/item/89t2v7h2>
- Xiang, Y., Vélez, N., & Gershman, S. J. (2023). Collaborative decision making is grounded in representations of other people’s competence and effort. *Journal of Experimental Psychology: General*, 152(6), 1565–1579. <https://doi.org/10.1037/xge0001336>
- Xiang, Y., Vélez, N., & Gershman, S. J. (2024). Optimizing competence in the service of collaboration. *Cognitive Psychology*, 150, Article 101653. <https://doi.org/10.1016/j.cogpsych.2024.101653>
- Zhang, X., McDougale, S. D., & Leonard, J. A. (2025). People accurately predict the shape but not the parameters of skill learning curves. *Cognition*, 258, Article 106083. <https://doi.org/10.1016/j.cognition.2025.106083>

## Appendix A

### Simulated Data Sets From Human-Fitted Models

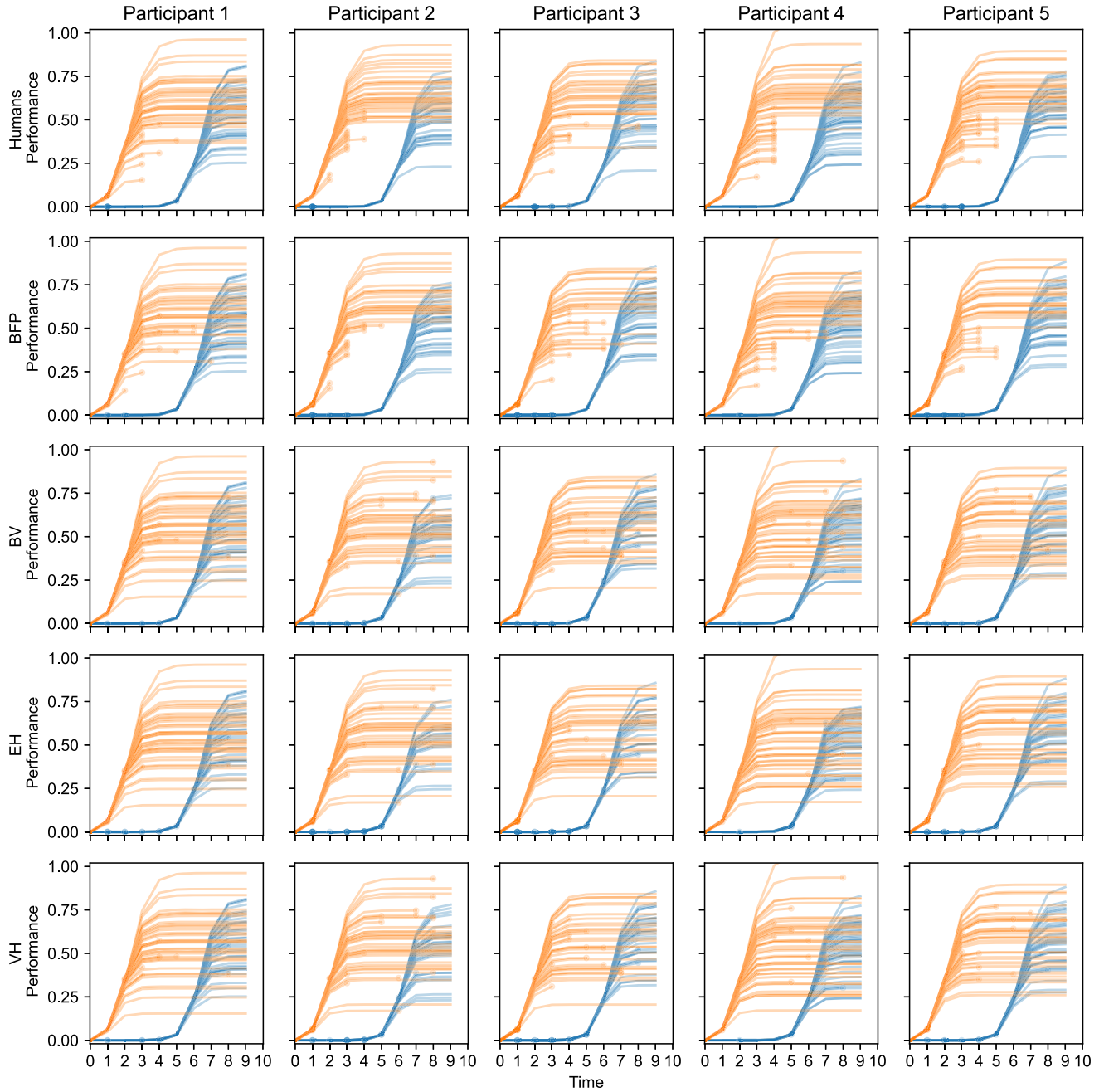
Experiments 3 and 4 exhibited rich patterns of participant behavior. To provide an analysis-agnostic characterization of raw human performance, individual variability, and model predictions, we generated simulated data sets from each fitted model and compared them directly with the empirical human data (Figures A1–A4). These simulated data sets serve as the foundation for subsequent model recovery analyses (Appendix B).

Among the candidate models, only the BFP model with logistic offset successfully reproduced key qualitative features

of human behavior. In Experiment 3, the BFP model, like participants, rarely switched in the separate-late condition and switched selectively when the starting asset’s asymptotic performance turned out to be below average in the separate-early condition. In Experiment 4, both humans and the BFP model exhibited condition-dependent switching, switching out at distinct timesteps across the peak-early and peak-late conditions.

*(Appendices continue)*



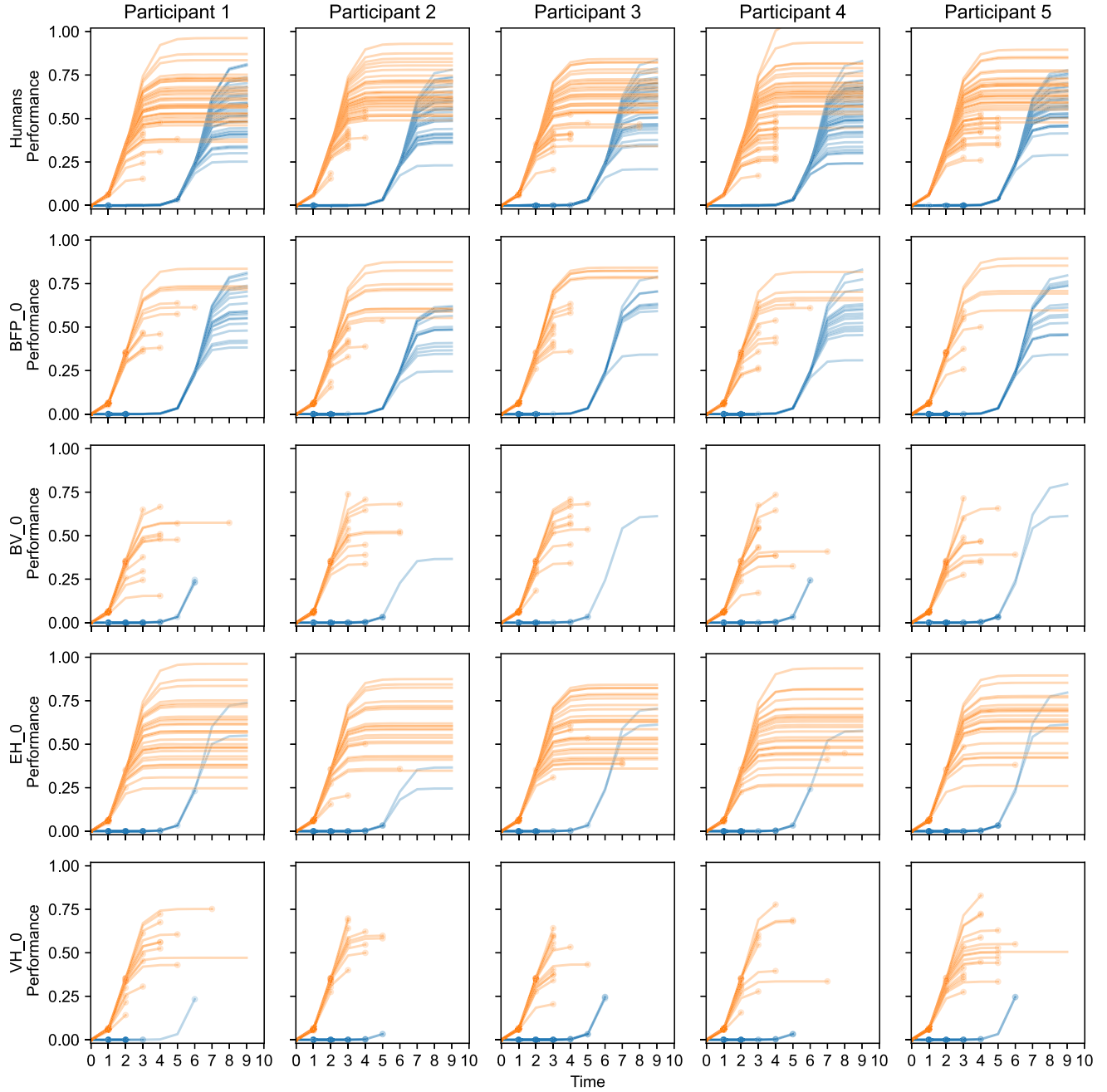
**Figure A1***Experiment 3 Data From Example Participants and Simulated Data From Each Fitted Model With Logistic Offset*

*Note.* Rows correspond to different fitted models (all with logistic offset  $b$ ). Different columns correspond to different participants. In each panel, the starting asset's performance for each task condition over time (colored lines) is visualized until the participant/simulated agent switches for that run (colored point). VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

*(Appendices continue)*

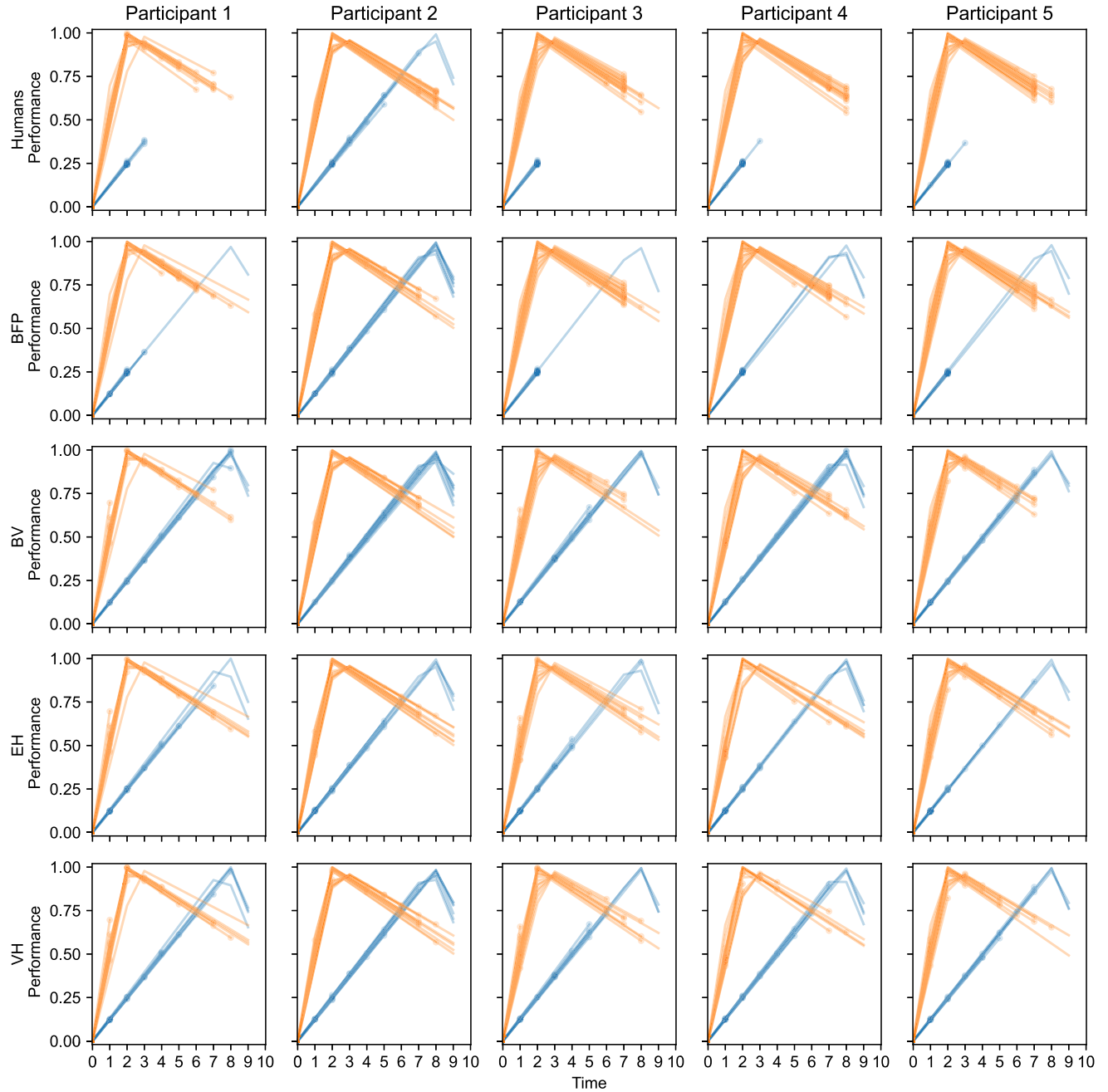
**Figure A2**

Experiment 3 Data From Example Participants and Simulated Data From Each Fitted Model Without Logistic Offset



*Note.* Rows correspond to different fitted models (all without logistic offset  $b$ ). Different columns correspond to different participants. In each panel, the starting asset's performance for each task condition over time (colored lines) is visualized until the participant/simulated agent switches for that run (colored point). VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

(Appendices continue)

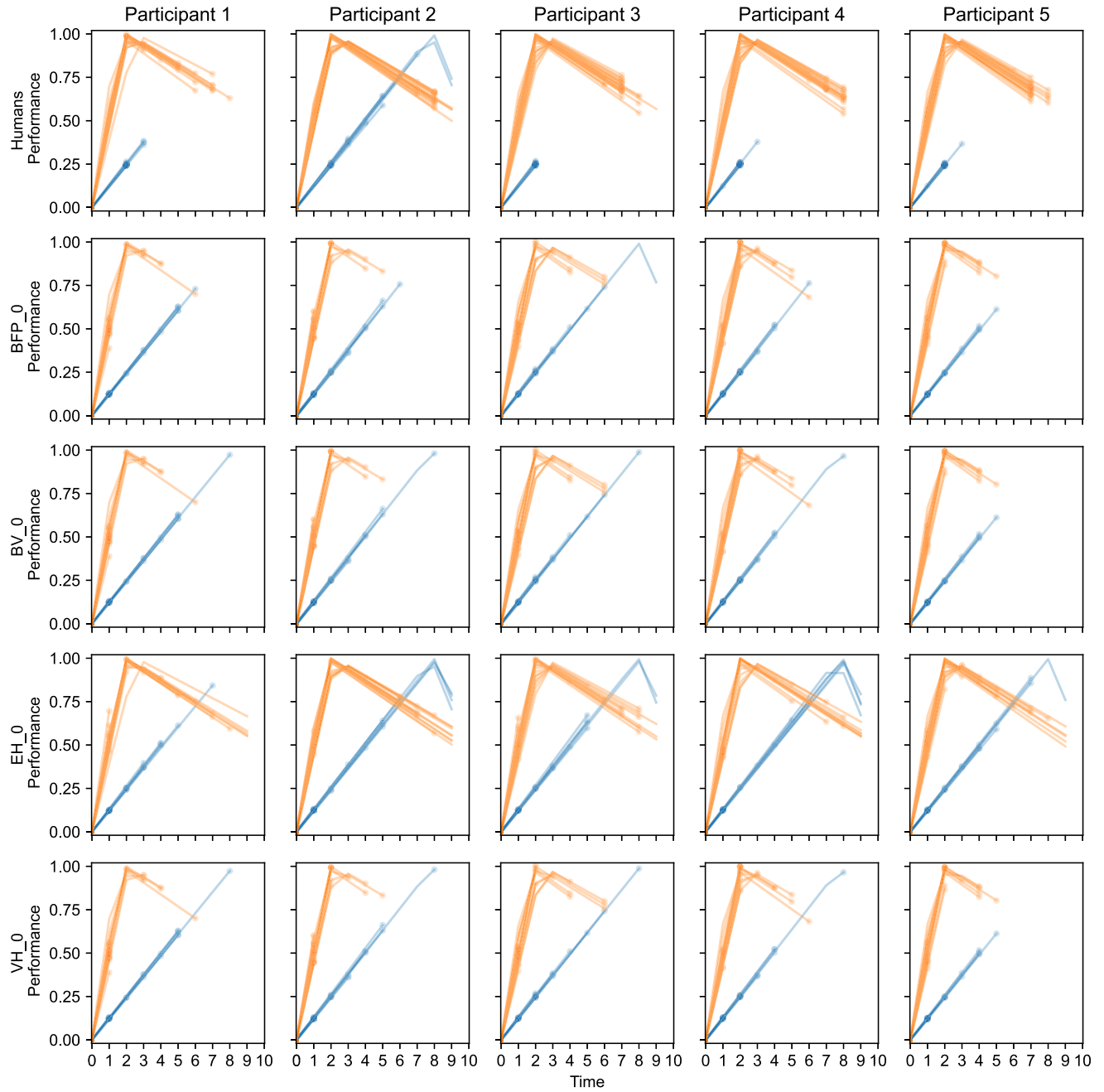
**Figure A3***Experiment 4 Data From Example Participants and Simulated Data From Each Fitted Model With Logistic Offset*

*Note.* Rows correspond to different fitted models (all with logistic offset  $b$ ). Different columns correspond to different participants. In each panel, the starting asset's performance for each task condition over time (colored lines) is visualized until the participant/simulated agent switches for that run (colored point). VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

*(Appendices continue)*

**Figure A4**

Experiment 4 Data From Example Participants and Simulated Data From Each Fitted Model Without Logistic Offset



*Note.* Rows correspond to different fitted models (all without logistic offset  $b$ ). Different columns correspond to different participants. In each panel, the starting asset's performance for each task condition over time (colored lines) is visualized until the participant/simulated agent switches for that run (colored point). VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

(Appendices continue)



## Appendix B

### Model Recovery

To evaluate whether our model-fitting procedures and fivefold cross-validation could identify the true generative model, we performed model recovery analyses for each experiment. For each candidate model, we generated simulated data sets using the parameter values fitted to human data (Appendix A) and then refitted all candidate models to each simulated data set using the same fitting and cross-validation procedures described in the main text. Finally, we examined whether the ground-truth model, refitted to its own simulated data set, featured the lowest test loss among all models fitted to the same data set.

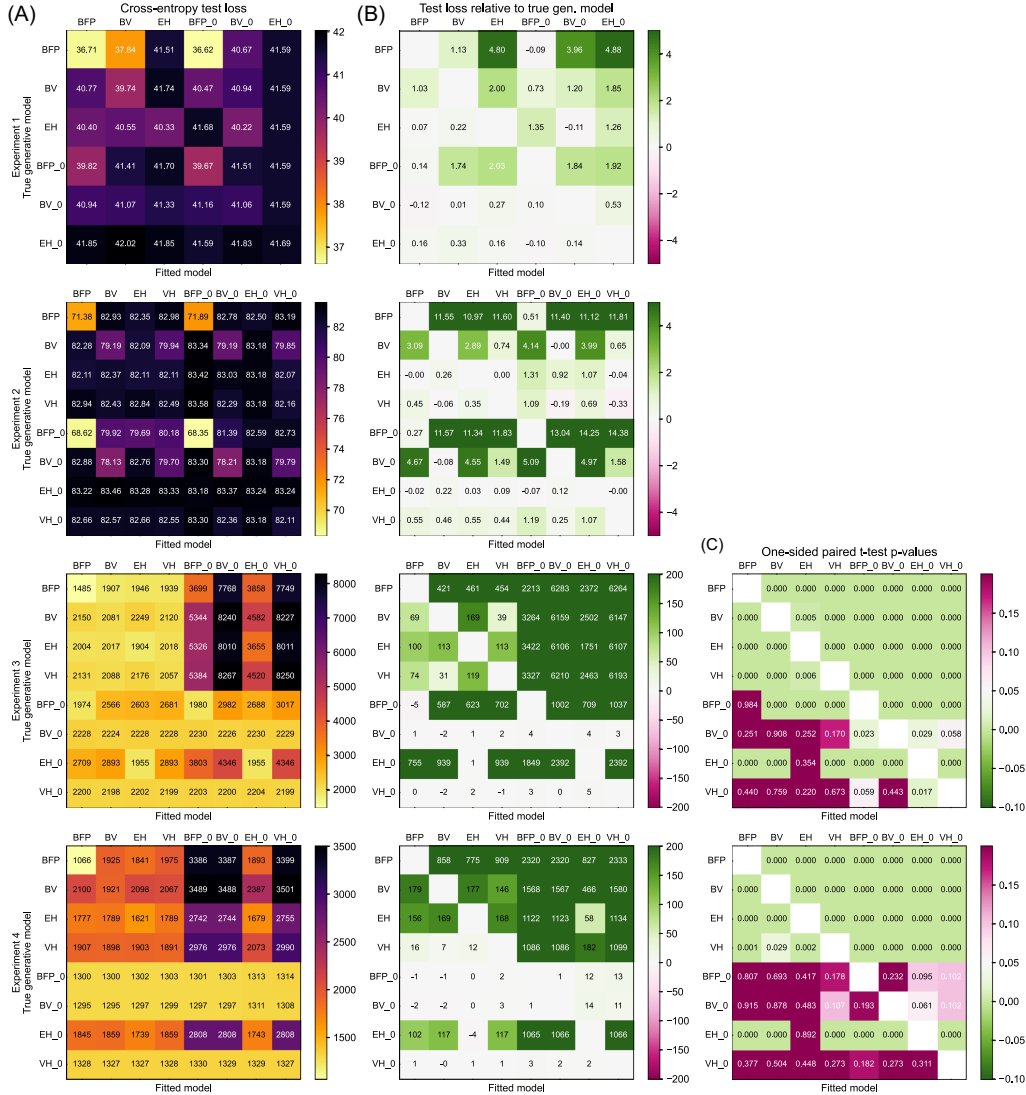
The results indicate that the modeling pipeline reliably recovers the ground-truth model whenever the generating model produces systematic behavior, even when this behavior includes stochastic variability (Figure B1). In contrast, data sets generated by models that produce fully random or nearly flat response probabilities cannot be uniquely recovered, which is theoretically expected. When a model produces choices insensitive to trial-specific information, any competing model can approximate its behavior by adopting an extremely flat logistic slope (i.e., high decision noise) and an appropriate offset. Under such conditions, the behavioral data contain no informative structure for model discrimination.

The above nuance primarily affects models that, when fitted to human data, yield near-identical response probabilities across timesteps or conditions (e.g., EH and EH\_0 in Experiments 1–2; BFP\_0, BV\_0, and VH\_0 in Experiment 4; Figures 2, 4, 10), whose simulated data sets also demonstrate such characteristics. Importantly, this degeneracy does not undermine the interpretability of model comparison on real human data. Human responses in all experiments show clear, structured variations across timesteps and conditions, meaning that identical-response models cannot plausibly account for their behavior. Similar reasoning applies to the BV\_0 model in Experiment 1: While the BFP model could capture data simulated from BV\_0, human switching probabilities are far from the near-flat patterns BV\_0 predicted (Figure 2A).

Excluding these degenerate cases, the recovery results confirm that our procedure successfully distinguishes the BFP, BV, EH, and VH models from one another. Critically, for Experiments 2–4, the BFP model with logistic offset (winning model comparison in the human fits) features the lowest cross-validated test loss only when it is the true generative model, and not otherwise. This supports the validity of the model comparison procedure used in the main text.

*(Appendices continue)*

**Figure B1**  
*Model Recovery Results for Each Experiment*



*Note.* Panel rows correspond to Experiments 1 to 4. (A) Participant-summed, fold-averaged test losses associated with fitting each model (columns) to the fake data set generated by each model under human-fitted parameter values (within-panel rows). (B) Difference in participant-summed fold-averaged test losses of different models to that of the true generated model. Green denotes higher losses (successful recovery), and red denotes lower losses. (C) One-sided  $t$  test  $p$  values over the individual-level fold-averaged test loss differences. One-sided tests were used to convey the direction of successful model recovery. Green denotes the significance of the multiplied  $p$  value at  $\alpha = .05$ , and red denotes nonsignificance. VH = velocity heuristic; EH = exploitation heuristic; BV = Bayesian velocity; BFP = Bayesian final performance. See the online article for the color version of this figure.

Received March 22, 2025  
 Revision received February 24, 2026  
 Accepted March 7, 2026 ■