A Rational Model of Growth Mindset Theory

Yuhao Michael Zhu (psymichaelzhu@gmail.com)[1] **& Bonan Zhao**[2]
[1]Social Science Division, University of Chicago
[2]School of Informatics, University of Edinburgh

Abstract

Mindset theory proposes that believing intelligence is improvable fosters academic achievement. Despite its influence, the theory lacks a mechanistic foundation. Here, we examine mindset-driven behaviors from the perspective of the computational problem the mind is solving—optimizing cumulative reward under uncertainty about skill malleability. We formalize this problem as a Markov decision process, where agents balance cultivating for future gains against harvesting immediate rewards. Growth- and fixed-mindsets are represented as the agent's priors over skill malleability. Through simulation, we demonstrate that mindset-driven behaviors, such as persistence and challenge seeking, arise as rational decisions under different prior beliefs, with optimistic priors promoting sustained engagement, belief updating, and higher rewards in favorable environments. Crucially, mindset effects vanish when environmental structures disincentivize long-term investment, leading agents to converge on a harvest-only strategy. Our model offers a unified account of the heterogeneous findings of mindset interventions, highlighting the importance of supportive environments for mindset effects to manifest.

Keywords: Bayesian model; growth mindset; skill learning; malleable environment; individual difference

Introduction

When learning a skill or training for a sport, sustained effort can incrementally enhance future payoffs by improving one's ability (Ericsson et al., 1993; Macnamara et al., 2014). However, investing in learning and training can be costly and may not necessarily pay off, creating a trade-off between continued investment and disengagement. Examining behavior in these environments is thus crucial for understanding human adaptability, and has attracted significant interest across learning, decision-making, and development (Denrell & Le Mens, 2020; Gureckis & Love, 2009; Hotaling et al., 2021; Nussenbaum & Hartley, 2024).

Among these, mindset theory is a particularly influential framework (Dweck, 2006; Dweck & Yeager, 2019). A mindset refers to an individual's beliefs about the malleability of human attributes such as intelligence. Those with a growth mindset believe that attributes can be improved through effort, while those with a fixed mindset see them as static. Consistent with theoretical predictions, growth-mindset individuals are more likely to persist (Hecht et al., 2022), embrace challenges (Yeager et al., 2019), and ultimately achieve higher academic performance (Jia et al., 2021). In contrast, people with a fixed mindset are more likely to give up and avoid challenges, leading to poor outcomes.

Despite its compelling insights, mindset theory remains largely descriptive. It links malleability beliefs to behaviors such as persistence and challenge seeking, but leaves open why these beliefs necessarily give rise to such behaviors and what decision-making processes underlie this relationship.

Moreover, empirical evaluations of mindset interventions have yielded heterogeneous results, suggesting that mindset effects might be context-dependent rather than universal (Canning et al., 2019; Jia et al., 2021; Yeager et al., 2019, 2022), for reviews, see (Hecht et al., 2023; Walton & Yeager, 2020; Yeager & Dweck, 2020). However, the original mindset theory offers limited specification of the boundary conditions that determine when these effects should appear.

In this study, we examine the computational problem implicit in mindset theory. We formalize mindset-driven behaviors as rational responses to inferred skill malleability within a sequential decision framework. Through simulation, we reproduce core predictions of mindset theory in terms of persistence and challenge seeking, and reveal the decision dynamics through which malleability beliefs shape behavior. We further show how environmental factors constrain—and in some cases eliminate—observable mindset effects.

Model

We formalize the computational problem underlying mindset theory as sequential decision-making under uncertainty: agents must trade off between exploiting their current skill level against investing in its improvement for future returns. We represent mindsets as priors over skill malleability, which shape utility estimates and resulting behaviors.

Malleable Skill

Definition We formalize the process of learning a skill as a Markov decision process (MDP). Each state represents a certain level of skill mastery. Skill level increases through practice, and higher levels translate into larger rewards. The agent therefore progresses through an ascending sequence of states $\{s_i\}_{i=0}^{\infty}$, corresponding to increasing skill levels.

Cultivate–harvest trade-off Over a finite horizon $T \in \mathbb{N}$, the agent balances harvesting short-term gains against cultivating for long-term benefits. At each time step t, the agent chooses between two actions: (1) *Cultivate* a_C, which will not provide immediate reward $r(s_i, a_C) = 0$ but may advance

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

the agent to the next state with probability $m \in (0, 1)$, which represents skill malleability, i.e., $p(s_{i+1} \mid s_i, a_C) = m$ and $p(s_i \mid s_i, a_C) = 1 - m$; (2) *Harvest* a_H, which provides an immediate payoff $r(s_i, a_H) = f(i)$ while leaving the state unchanged, i.e., $p(s_i \mid s_i, a_H) = 1$. Intuitively, cultivation corresponds to investing effort in skill improvement, such as learning a new guitar scale, whereas harvesting corresponds to exploiting existing skills for immediate gratification, such as playing familiar scales.

We use a linear reward function $f : \mathbb{N} \to \mathbb{R}$ to denote the mapping from state index i to harvest payoff

$$f(i) = R + w\, i, \tag{1}$$

where $w > 0$ is reward growth rate and R is reward baseline.

Optimal switch point According to Zhao et al. (2024), this decision problem corresponds to an optimal stopping problem (Chow et al., 1971): the agent should cultivate up to a switch point and then exclusively harvest for the remaining steps.

Suppose the agent is at time step t with current skill level I_t. Let $D = T - t$ denote the number of remaining steps, and let $R_t = R + wI_t$ denote the current harvest payoff. Under a candidate stopping policy $\pi(d)$, the agent cultivates for the next $d \in \{0, \ldots, D\}$ steps and then harvests for the remaining $D - d$ steps. If the true malleability is m, then the number of successful improvements during the d cultivation attempts is $X_d \sim \text{Binomial}(d, m)$. The expected return under $\pi(d)$ is

$$\begin{aligned} G_t(d) &= (D - d)\, \mathbb{E}[f(I_t + X_d)] \\ &= (D - d)\, (R_t + wmd)\,. \end{aligned} \tag{2}$$

Since $G_t(d)$ is a concave-down quadratic function of d ($w > 0$ and $m > 0$), it has a maximum achieved at

$$\tilde{d}_t = \frac{D}{2} - \frac{R_t}{2wm}. \tag{3}$$

The optimal switch point is thus the closest feasible integer

$$d_t^* = \text{clip}_{[0,D]} \left(\text{round} \left[\frac{D}{2} - \frac{R_t}{2wm} \right] \right). \tag{4}$$

Importantly, d_t^* is recomputed at every step based on the agent's current skill level I_t and remaining horizon $D = T - t$.

Skill with Unknown Malleability

The above formalization assumes that the agent knows the true skill malleability m. In reality, however, the agent rarely knows in advance how hard it is to improve the skill. For example, when learning guitar, one might not know whether it will take hours or days of practice to play a piece fluently.

Malleability belief We model the agent's prior over skill malleability using a Beta distribution: $b_0 \sim \text{Beta}(\alpha, \beta)$, where the prior mean $\hat{b}_0 = \alpha/(\alpha + \beta)$ is the agent's initial estimate of skill malleability (Fig. 1).

After each cultivation attempt, the agent observes whether the attempt succeeds. After observing x_t successes out of n_t cultivation attempts, the agent updates their posterior belief according to

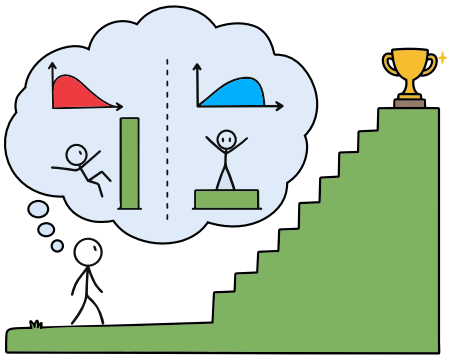

Figure 1: Malleability belief. The agent decides whether to attempt to move up to a higher skill level (*Cultivate*), or to remain in the comfort zone (*Harvest*). A growth-mindset agent, shown in blue, holds an optimistic prior that cultivation will succeed. In contrast, a fixed-mindset agent, shown in red, holds a pessimistic prior of cultivation success.

$$b_t \sim \text{Beta}(\alpha + x_t,\ \beta + n_t - x_t). \tag{5}$$

The posterior mean $\hat{b}_t = (\alpha + x_t)/(\alpha + \beta + n_t)$ serves as the agent's current estimate of skill malleability. There is no belief updating for the harvest action.

Cultivate-harvest policy At each time step t, the agent substitutes their current estimate $\hat{b}_t$ for the unknown malleability m, and recomputes the optimal switch point $\hat{d}_t^*$ as in Eq. 4, and chooses an action according to

$$a_t = \begin{cases} a_C, & \text{if } \hat{d}_t^* > 0, \\ a_H, & \text{if } \hat{d}_t^* = 0. \end{cases} \tag{6}$$

Eq. 6 states that a rational agent should cultivate if the optimal switch point has not been reached, or harvest otherwise. Eq. 4 further indicates that a higher estimate of skill malleability $\hat{b}_t$, a higher reward growth rate w, or a longer remaining horizon $D = T - t$ encourages longer cultivation, whereas a larger harvest reward $R_t = R + wI_t$ may shorten cultivation.

Choosing Among Skills

So far, we have derived the optimal policy in situations with only one skill, where the agent decides whether to cultivate or harvest that skill. This formalization can also be extended to situations involving multiple skills.

Multiple skills The agent chooses among skills over a time horizon T. Each skill is represented as an MDP as above, and we use the subscript τ to denote the skill type. Each skill has its own parameters: malleability m_τ, reward baseline R_τ, and reward growth rate w_τ. Similarly, the agent's prior about skill τ's malleability is modeled by $b_{0,\tau} \sim \text{Beta}(\alpha_\tau, \beta_\tau)$.

Skill selection policy At each time step t, the agent estimates the value of each skill based on current skill level and malleability belief. Let $I_{t,\tau}$ denote the agent's current level in skill τ, and define the current harvest payoff as $R_{t,\tau} = R_\tau + w_\tau I_{t,\tau}$. A rational agent should estimate the optimal switch point for each skill:

$$\hat{d}^*_{t,\tau} = \text{clip}_{[0,D]}\left(\text{round}\left[\frac{D}{2} - \frac{R_{t,\tau}}{2w_\tau \hat{b}_{t,\tau}}\right]\right). \quad (7)$$

The corresponding optimal expected return is thus

$$\hat{G}^*_{t,\tau} = (D - \hat{d}^*_{t,\tau})\left(R_{t,\tau} + w_\tau \hat{b}_{t,\tau}\hat{d}^*_{t,\tau}\right). \quad (8)$$

The agent then selects the skill with the highest return:

$$\tau_t = \arg\max_\tau \hat{G}^*_{t,\tau}. \quad (9)$$

After choosing skill τ_t, the agent cultivates that skill if $\hat{d}^*_{t,\tau_t} > 0$ and harvests otherwise. If cultivation is attempted, the belief about the chosen skill is updated according to Eq. 5; beliefs about unchosen skills remain unchanged.

Simulation

Even when agents follow the same optimal policy in the same environment, their behaviors can diverge due to differing prior beliefs about skill malleability—corresponding to mindsets in mindset theory (Dweck & Yeager, 2019)—which shape all subsequent behaviors. To characterize the role of prior differences, we conducted simulations within a unified skill development task, systematically manipulating agents' malleability priors and environmental parameters. The code and data for simulations are available at https://github.com/psymichaelzhu/CogSci_mindset.

Skill Development Task

Setup We design a task that mimics real-world skill acquisition, in which agents repeatedly decide what to learn and how to engage, balancing immediate reward harvesting against long-term skill cultivation.

There are two skills: an *Easy* skill and a *Challenging* skill. In each trial, the agent first selects one skill and then decides how to engage with it: *Cultivate*, to potentially improve future performance, or *Harvest*, to exploit the current level of mastery for immediate rewards.

Each skill operates as an independent Markov decision process. Two skills differ in their reward structures. Compared to the *Easy* skill, the *Challenging* skill has a lower reward baseline ($R_{\text{Challenging}} = R_{\text{Easy}}/2$) but a higher reward growth rate ($w_{\text{Challenging}} = 2 \cdot w_{\text{Easy}}$). This contrast reflects the idea that easy skills provide instant gratification yet quickly become boring, while challenging skills initially frustrate but ultimately lead to a deeper sense of achievement.

We set the two skills to share the same malleability, $m_{\text{Easy}} = m_{\text{Challenging}} = m$. Although skills may also differ in malleability in real-world settings, this simplification keeps our setup parsimonious and isolates the effect of malleability beliefs from objective differences in skill malleability.

Decision rule The rational agent updates their beliefs about each skill's malleability independently (Eq. 5). At each time step, the agent estimates the optimal switch point for each skill based on the current belief (Eq. 7) and computes the corresponding expected return (Eq. 8). The agent then selects the skill with the higher return (Eq. 9). Once a skill is chosen, the agent follows the optimal cultivate–harvest policy: cultivate if the current trial has not yet reached the skill's optimal switch point or harvest otherwise (Eq. 6).

This two-stage decision repeats at each time step. When the agent chooses a skill and an action, only the chosen skill's level may be updated. A successful cultivation attempt increases the chosen skill's level by one, whereas a failed cultivation attempt or harvesting leaves its level unchanged; unchosen skills remain unchanged. After observing the cultivation outcome, the agent updates their belief about the chosen skill's malleability, whereas unchosen skills and harvesting action provide no information.

As both skill levels and malleability beliefs evolve, the agent may revise their estimates of each skill's expected return and switch point, which may consequently cause shifts in behaviors.

Parameterization

Mindsets as malleability priors We operationalize mindsets as prior beliefs about skill malleability and vary these priors along two independent dimensions: (1) *prior estimate*, defined as the mean of the prior distribution, $\hat{b}_0 = \alpha/(\alpha+\beta)$, ranges from 0.1 to 0.9 in increments of 0.1, capturing a continuum from fixed- to growth-oriented mindsets. (2) *prior strength*, defined as the concentration of the prior distribution, $\kappa_0 = \alpha + \beta$, ranges from 10 to 60 in increments of 10, capturing increasingly strong commitments to prior beliefs.

Environmental parameters We manipulate four environmental parameters: (1) *skill malleability* $m \in \{0.3, 0.5, 0.7\}$, higher values indicate that skill level is more likely to increase through cultivation; (2) *reward baseline* $R \in \{3, 51, 99\}$, higher values indicate greater baselines for harvest rewards; (3) *reward growth rate* $w \in \{0.9, 7.5, 14.1\}$, higher values indicate that harvest rewards increase more steeply as skill level improves; (4) *time horizon* $T \in \{10, 45, 80\}$, higher values indicate more available time steps.

Reward baseline R and reward growth rate w are mapped onto the two skills to create contrasting reward profiles. The *Easy* skill has a higher reward baseline but flatter reward growth ($R_{\text{Easy}} = \frac{4}{3}R$, $w_{\text{Easy}} = \frac{2}{3}w$), whereas the *Challenging* skill has a lower reward baseline but steeper reward growth ($R_{\text{Challenging}} = \frac{2}{3}R$, $w_{\text{Challenging}} = \frac{4}{3}w$). The two skills share the same malleability $m_{\text{Easy}} = m_{\text{Challenging}} = m$.

Measures

We examine two core behavioral outcomes described by mindset theory: (1) *challenge seeking*, defined as the proportion of steps in which the agent engages with the *Challenging* skill, either by cultivating or harvesting; (2) *persistence*, defined as

the overall proportion of steps spent cultivating before switching to harvesting, averaged across skills.

In addition to these behavioral outcomes, we evaluate task performance and belief adaptation: (3) *cumulative reward* is defined as the total reward obtained over the task; (4) *belief updating* is defined as the change in the agent's malleability estimate over the course of the task, averaged across skills. Updating beliefs toward the true malleability improves the accuracy of the agent's internal model, thereby supporting future adaptability.

Results

We test the theoretical predictions by examining behavioral outcomes associated with mindset and further investigate the underlying decision-making mechanisms. In addition, this formalization allows us to identify boundary conditions of mindset effects by examining when environmental factors produce or eliminate observable behavioral differences.

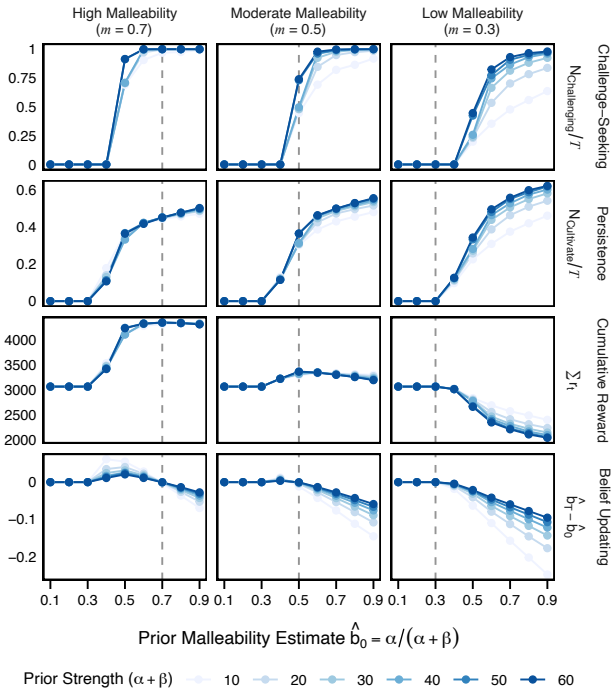

Figure 2: Behavioral outcomes of different mindsets. Mindsets are operationalized as priors over skill malleability. Points represent averages over 2000 simulation runs. Vertical dashed lines indicate the true malleability shared by both skills. Belief updating (last row) is averaged across skills. All environmental parameters other than malleability are held at moderate levels: $R = 51$, $w = 7.5$, and $T = 45$.

Reproducing Predictions of Mindset Theory

We observe relatively stable mindset effects in our simulations. As shown in Fig. 2, when malleability priors shift toward the growth end (larger values on the x-axis), the agent engages more with the *Challenging* skill (first row) and cultivates for longer periods (second row). These patterns capture the key behavioral predictions of mindset theory: individuals who believe that abilities are improvable exhibit greater challenge seeking and longer persistence.

Interestingly, a growth mindset does not always correspond to better task performance (Fig. 2, third row). When skill malleability is inherently low, growth-oriented agents obtain lower cumulative reward (right column). In contrast, when skills are sufficiently malleable (left and middle columns), greater challenge seeking and persistence translate into higher cumulative reward. Together, these results suggest that the adaptivity of mindset-driven behaviors depends on the true malleability of skills.

Our simulation further reveals an asymmetry in belief updating. Agents with growth-oriented priors revise their beliefs more strongly toward the true skill malleability (Fig. 2, last row). This belief correction is most pronounced when prior strength is weak (lighter-colored lines). Importantly, this asymmetry is not simply driven by prior accuracy. For skills with moderate malleability (middle column), even when growth- and fixed-mindset agents are initialized at equal distances from the true malleability, only growth-mindset agents substantially update their beliefs, whereas fixed-mindset agents show little correction. Such correction may allow growth-mindset agents to adapt more effectively to the environment over time, thereby partially offsetting their reward losses in low-malleability settings.

Mechanism underlying Mindset Effects

Overall, malleability beliefs affect behavior through two related components: (1) *policy-based computations*: for each skill, the agent estimates the optimal switch point (Eq. 7) and the corresponding expected return (Eq. 8) based on their current belief; (2) *choice implementation*: according to these estimates, the agent chooses which skill to engage with (*Easy* vs. *Challenging*, Eq. 9) and how to engage with it (*Cultivate* vs. *Harvest*, Eq. 6).

As agents' malleability priors shift from fixed- to growth-oriented, they estimate higher expected returns for both the *Easy* and *Challenging* skills, with a greater increase for the *Challenging* skill (Fig. 3A). This produces a reversal in skill preference along the mindset continuum: fixed-oriented agents assign relatively higher value to the *Easy* skill, whereas growth-oriented agents assign relatively higher value to the *Challenging* skill. This reversal explains the increase in challenge-seeking behavior as priors become more growth-oriented (Fig. 2, top row). Similarly, higher malleability priors lead agents to plan longer periods of cultivation before switching to harvesting (Fig. 3B), accounting for the increased persistence observed earlier (Fig. 2, second row).

These computational differences translate into distinct behavioral trajectories. Agents with low malleability priors ($\hat{b}_0 = 0.3$) expect relatively higher return from the *Easy* skill, which is paired with an optimal switch point at zero (left dashed line in Fig. 3A–B). Therefore, they immediately

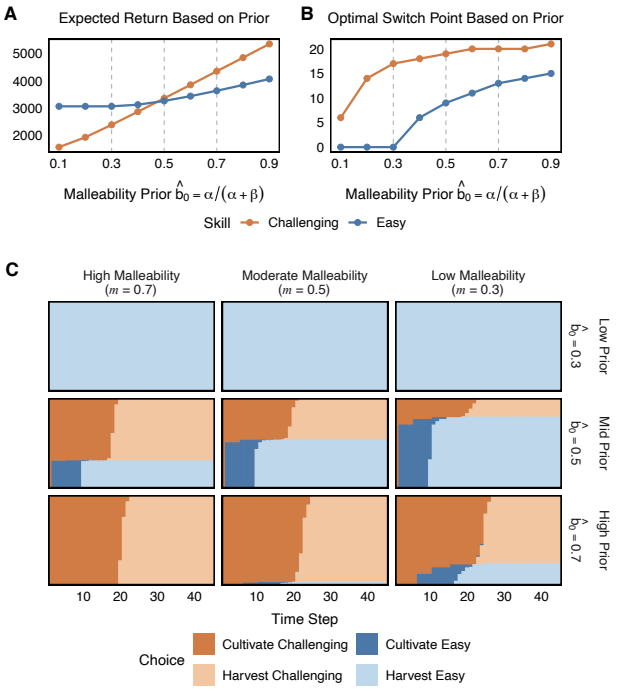

Figure 3: Mechanisms underlying mindset effects. (A) Expected returns and (B) optimal switch points for each skill, calculated based on agents' malleability priors. (C) Choice dynamics across 2000 simulations. The y-axis shows, for each trial, the proportion of agents selecting a certain combination of skill (*Easy* vs. *Challenging*) and action (*Cultivate* vs. *Harvest*). Environmental parameters are $R = 51$, $w = 7.5$, and $T = 45$. Prior strength is fixed at 30.

harvest the *Easy* skill without cultivation (Fig. 3C, top row). This strategy generates no feedback about skill malleability, thereby preventing belief updating (Fig. 2, bottom row). These agents perform well in low-malleability environments, where harvesting the *Easy* skill is effective, but perform poorly in moderate- and high-malleability environments, where active cultivation would be more beneficial (Fig. 2, third row).

Agents with high malleability priors ($\hat{b}_0 = 0.7$) show the opposite pattern. They expect higher return from the *Challenging* skill, with a substantially large optimal switch point (right dashed line in Fig. 3A–B). As a result, they start by cultivating the *Challenging* skill and are less likely to abandon it in response to early fluctuations in outcomes (Fig. 3C). This sustained engagement provides more feedback about skill malleability, leading to stronger belief updating (Fig. 2, bottom row). The sustained engagement also yields high payoffs in high-malleability environments, where cultivation is effective, but may impair payoffs in low-malleability environments, where cultivating the *Challenging* skill is less adaptive (Fig. 2, third row). Nevertheless, active engagement gives these agents more information for future adaptation.

Agents with intermediate malleability priors ($\hat{b}_0 = 0.5$) fall between these two extremes. They assign slightly higher value to the *Challenging* skill than to the *Easy* skill and estimate nonzero optimal switch points for both skills (middle dashed line in Fig. 3A–B). Thus, they also initially cultivate the *Challenging* skill, but are more likely to shift toward the *Easy* skill as feedback is realized over time (Fig. 3C, middle row). This flexible strategy produces moderate belief updating (Fig. 2, bottom row) and stabilizes performance in low-malleability environments (Fig. 2, third row).

Boundary Conditions of Mindset Effects

We manipulate environmental parameters to identify the conditions under which mindset effects are likely to emerge. We find that mindset effects are strongly constrained by temporal and incentive structures (Fig. 4). Pronounced behavioral differences emerge only within an intermediate range of environmental parameters.

When the environment discourages long-term investment (right side of each panel)—through reduced incentives for improvement (lower reward growth rates), increased time pressure (shorter time horizons), or greater temptation for instant gratification (higher reward baselines)—agents with different mindsets converge on harvesting the *Easy* skill.

At the opposite extreme, when the environment strongly encourages long-term investment (left side of each panel)—characterized by high reward growth rates, long time horizons, or low reward baselines—agents again behave similarly despite differing priors. In these environments, both growth- and fixed-mindset agents actively cultivate the *Challenging* skill, resulting in indistinguishable behavioral outcomes.

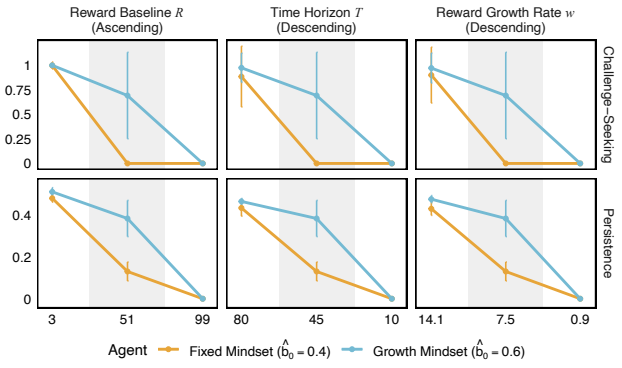

Figure 4: Boundary conditions of mindset effects. Growth-mindset and fixed-mindset agents are assigned Beta$(6, 4)$ and Beta$(4, 6)$ priors, respectively, with skill malleability fixed at $m = 0.5$, equidistant from both prior means. In each panel, only one environmental parameter is varied; the remaining parameters are held at moderate values ($R = 51$, $w = 7.5$, $T = 45$), indicated by the light gray regions. The x-axes for time horizon and reward growth rate are inverted. Points denote means and error bars denote standard deviations across 2000 simulations.

Discussion

Mindset theory has been an influential theoretical framework in psychology. In this paper, we aim to provide a normative account of mindset theory by formalizing persistence and challenge seeking as rational responses to inferred skill malleability in a Bayesian sequential decision-making framework.

Our simulations reproduce the core predictions of mindset theory: growth-oriented agents show greater challenge seeking and longer persistence, which in turn promote belief updating and yield higher payoffs in favorable environments. We further show that these effects arise because malleability priors alter agents' estimated optimal switch points and expected returns, thereby shaping the dynamics of skill selection and cultivate–harvest decisions. Finally, we identify boundary conditions for mindset effects: mindset-driven differences emerge under moderate environmental parameters. When long-term investment is either strongly discouraged or encouraged, agents with different priors converge on similar strategies, eliminating observable behavioral differences.

Our analysis provides insight into the increasingly emphasized context-dependence of mindset effects (Hecht et al., 2023; Walton & Yeager, 2020). Empirical studies of mindset interventions have produced heterogeneous findings, and existing accounts have often attributed this heterogeneity to social factors such as peer norms (Yeager et al., 2019) or teacher mindset beliefs (Canning et al., 2019; Yeager et al., 2022). Our simulations complement this perspective with the following insights. First, agents with different mindsets may behave similarly when long-term investment is either strongly discouraged or encouraged (Fig. 4). This suggests that even if an intervention successfully changes individuals' mindsets, the behavioral consequences may remain invisible when the measurement context is poorly suited to reveal mindset-driven differences. Second, the belief-updating results in Fig. 2 suggest that strong priors limit the extent to which agents update from experience. Moreover, fixed-oriented priors may prevent agents from engaging in cultivation, thereby reducing the feedback needed for belief revision. This resembles the learning trap effect (Rich & Gureckis, 2018), in which choice-contingent feedback prevents agents from gaining information that would correct their beliefs. Finally, the payoff results in Fig. 2 show that growth mindset is not always adaptive (Dweck & Yeager, 2019): when a skill is in fact difficult to improve, growth-mindset agents' active engagement can instead lead to worse performance.

Broadly speaking, our model illustrates the value of formalization (Cushman, 2024). It translates verbally specified assumptions into explicit computational components, allowing quantitative simulation and opening space for future extensions. First, the model could be extended to capture how mindsets generalize across domains. Individuals may hold malleability beliefs about different domains, and these beliefs may be either relatively independent or shaped by a higher-level general mindset. Such a structure could be naturally formalized with a hierarchical Bayesian model. Second, the framework could be used to prospectively evaluate intervention strategies. Because educational interventions are often costly to implement, simulations provide a way to evaluate their potential effects before empirical testing. For example, intelligence praise and effort praise (Mueller & Dweck, 1998) could be formalized as different feedback algorithms and compared in terms of their effects on belief updating. Third, the model could be extended to multi-agent settings. Individuals' mindsets are embedded in social environments and shaped by others' norms and feedback. Future models could consider interactions among multiple agents to examine how malleability beliefs propagate across individuals. Fourth, symbolic abstraction helps reveal connections between mindset theory and broader domains beyond education, such as innovation behavior (Zhao et al., 2024). Finally, beyond theoretical extensions, our skill development task could also be developed into an empirical behavioral paradigm, providing a complementary measure of mindset beyond self-report. The task might also serve as a potential intervention tool: alongside existing message-based approaches, a game-like paradigm could convey growth-oriented ideas through direct experience.

The present analysis also has its limitations. First, we simplify cultivate and harvest as opposing actions, but in reality these two processes are not always mutually exclusive. For example, playing a song can both showcase current mastery and strengthen future performance. This simplification is nevertheless useful, because it highlights the core computational tension in mindset theory. Second, we assume that agents maximize cumulative reward, although actual behaviors may also be influenced by factors such as intrinsic reward and social evaluation. Our goal is not to deny these factors, but to show that even a minimal rational model can generate systematic behavioral differences from different malleability priors. Future models could incorporate these additional motives into the utility function. Finally, our task allows agents to choose which skill to pursue, whereas such choices may be constrained in educational settings. That said, even when students cannot freely choose their subjects, they can still allocate effort across subjects and decide when to persist or disengage. Moreover, because the core cultivate–harvest trade-off arises in the single-skill version of the model, our framework does not depend on unconstrained skill selection.

In sum, by formalizing mindset theory within a normative decision-making model, this work shows how malleability priors and environmental structures jointly shape adaptive engagement in skill learning. We hope this framework can inform future interventions and offer a computational perspective on how beliefs guide behavior in malleable environments.

References

Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, *5*(2), eaau4734. https://doi.org/10.1126/sciadv.aau4734

Chow, Y. S., Robbins, H., & Siegmund, D. (1971). *Great expectations: The theory of optimal stopping*.

Cushman, F. (2024). Computational social psychology. *Annual Review of Psychology*, *75*, 625–652. https://doi.org/10.1146/annurev-psych-021323-040420

Denrell, J., & Le Mens, G. (2020). Revisiting the competency trap. *Industrial and Corporate Change*, *29*(1), 183–205.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Dweck, C. S., & Yeager, D. S. (2019). Mindsets: A view from two eras. *Perspectives on Psychological Science*, *14*(3), 481–496. https://doi.org/10.1177/1745691618804166

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Gureckis, T. M., & Love, B. C. (2009). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition*, *113*(3), 293–313. https://doi.org/10.1016/j.cognition.2009.03.013

Hecht, C. A., Latham, A. G., Buskirk, R. E., Hansen, D. R., & Yeager, D. S. (2022). Peer-modeled mindsets: An approach to customizing life sciences studying interventions. *CBE—Life Sciences Education*, *21*(4), ar82. https://doi.org/10.1187/cbe.22-07-0143

Hecht, C. A., Murphy, M. C., Dweck, C. S., Bryan, C. J., Trzesniewski, K. H., Medrano, F. N., Giani, M., Mhatre, P., & Yeager, D. S. (2023). Shifting the mindset culture to address global educational disparities. *npj Science of Learning*, *8*(1), 29. https://doi.org/10.1038/s41539-023-00181-y

Hotaling, J. M., Navarro, D. J., & Newell, B. R. (2021). Skilled bandits: Learning to choose in a reactive world. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(6), 879–905. https://doi.org/10.1037/xlm0000981

Jia, L., Lim, C. H., Ismail, I., & Tan, Y. C. (2021). Stunted upward mobility in a learning environment reduces the academic benefits of growth mindsets. *Proceedings of the National Academy of Sciences*, *118*(10), e2011832118. https://doi.org/10.1073/pnas.2011832118

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, *25*(8), 1608–1618. https://doi.org/10.1177/0956797614535810

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Nussenbaum, K., & Hartley, C. A. (2024). Understanding the development of reward learning through the lens of meta-learning. *Nature Reviews Psychology*, *3*(6), 424–438. https://doi.org/10.1038/s44159-024-00304-1

Rich, A. S., & Gureckis, T. M. (2018). The limits of learning: Exploration, generalization, and the development of learning traps. *Journal of Experimental Psychology: General*, *147*(11), 1553–1570. https://doi.org/10.1037/xge0000466

Walton, G. M., & Yeager, D. S. (2020). Seed and soil: Psychological affordances in contexts help to explain where wise interventions succeed or fail. *Current Directions in Psychological Science*, *29*(3), 219–226. https://doi.org/10.1177/0963721420904453

Yeager, D. S., Carroll, J. M., Buontempo, J., Cimpian, A., Woody, S., Crosnoe, R., Muller, C., Murray, J., Mhatre, P., Kersting, N., Hulleman, C., Kudym, M., Murphy, M., Duckworth, A. L., Walton, G. M., & Dweck, C. S. (2022). Teacher mindsets help explain where a growth-mindset intervention does and doesn't work. *Psychological Science*, *33*(1), 18–32. https://doi.org/10.1177/09567976211028984

Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist*, *75*(9), 1269–1284. https://doi.org/10.1037/amp0000794

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., ... Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*(7774), 364–369. https://doi.org/10.1038/s41586-019-1466-y

Zhao, B., Vélez, N., & Griffiths, T. (2024). A rational model of innovation by recombination. *Proceedings of the 46th Annual Meeting of the Cognitive Science Society*.