



Make-or-Break: Chasing Risky Goals or Settling for Safe Rewards?

Pantelis P. Analytis,^a Charley M. Wu,^b Alexandros Gelastopoulos^c

^aDanish Institute of Advanced Study, University of Southern Denmark

^bCenter for Adaptive Rationality, Max Planck Institute for Human Development

^cDepartment of Mathematics and Statistics, Boston University

Received 31 May 2018; received in revised form 14 April 2019; accepted 23 April 2019

Abstract

Humans regularly pursue activities characterized by dramatic success or failure outcomes where, critically, the chances of success depend on the time invested working toward it. How should people allocate time between such *make-or-break* challenges and safe alternatives, where rewards are more predictable (e.g., linear) functions of performance? We present a formal framework for studying time allocation between these two types of activities, and we explore optimal behavior in both one-shot and dynamic versions of the problem. In the one-shot version, we illustrate striking discontinuities in the optimal time allocation policy as we gradually change the parameters of the decision-making problem. In the dynamic version, we formulate the optimal strategy—defined by a *giving-up* threshold—which adaptively dictates when people should stop pursuing the make-or-break goal. We then show that this strategy is computationally inaccessible for humans, and we explore boundedly rational alternatives. We compare the performance of the optimal model against (a) a myopic giving-up threshold that is easier to compute, and even simpler heuristic strategies that either (b) only decide whether or not to start pursuing the goal and never give up or (c) consider giving up at a fixed number of control points. Comparing strategies across environments, we investigate the cost and behavioral implications of sidestepping the computational burden of full rationality.

Keywords: Resource allocation; Dynamic decision-making; Uncertainty; Risky choice; Perseverance; Sigmoid curves; Bounded rationality

Life can only be understood backwards, but it must be lived forwards

— Søren Kierkegaard.

1. Introduction

From applying for a research grant to giving one's all to earn a bonus at work, numerous human activities can either be crowned by dramatic success or thwarted by failure. In these "make-or-break" endeavors, people are either handsomely rewarded upon success or gain nothing upon failure. Although there is almost always a monotonic relationship between the invested time (or effort) in an activity and the expected levels of performance, there is considerable uncertainty about the outcomes (e.g., success or failure). Thus, some of the most important decisions in life may be defined by how one chooses to allocate limited resources between make-or-break tasks—with potentially life changing outcomes—and "safe" alternatives, where outcomes are a more predictable function of performance (e.g., the linear relationship a wage worker experiences between hours worked and income). How should you divide your time between chasing a challenging goal and investing in safer occupations? What is the optimal allocation strategy, and how does it compare to simpler, but computationally less expensive strategies? What behaviors do the different strategies produce across decision-making settings?

1.1. Background

As early as the 1930s, Kurt Lewin and colleagues identified the importance of success-or-failure reward structures for understanding human behavior (Hoppe, 1931; Lewin, 1936). They investigated how people with different abilities set their aspiration levels (Festinger, 1942; Rotter, 1942), calibrated based on the chances of success and the potential rewards to be gained. Lewin and colleagues postulated that people experience joy when they achieve their aspirations and feel acute loss when they fail, and that they choose the appropriate aspiration level simply by balancing these two motivating forces and the probability of success (Lewin, Dembo, Festinger, & Sears, 1944; Siegel, 1957). In their work, they assumed that people strictly commit themselves to a single aspiration level selected from a continuum of possible choices (i.e., a single task) and that there are no constraints on the amount of time or effort that could be invested (i.e., an infinite horizon). This early work provided valuable insights into how the likelihood of success influences how people set aspiration levels. Yet how do these insights transfer to problems where there is a finite horizon (i.e., limited time to be allocated) and multiple tasks with different payoff functions?

Since then, there have been repeated attempts in the behavioral sciences—independent of or inspired by Lewin's framework—to address how people reason about their chances of success in meeting their goals (e.g., Bandura, 1977; Heider, 1958; Weiner & Kukla, 1970), how aspiration levels affect people's choices (Diecidue & Van De Ven, 2008; Payne, Laughhunn, & Crum, 1980; Simon, 1955), and how people choose among variable levels of aspiration (Atkinson, 1957). Yet there has been no rigorous mathematical formalization of the problem people grapple with when they have to allocate scarce resources between a make-or-break activity, where investment of additional effort can

alter their chances of success, and a safe alternative.¹ The closest attempt to formally ground the problem was a static effort allocation framework developed by Kukla (1972), who greatly simplified the problem by assuming away any uncertainty in the relation between effort and performance. Kukla revealed a distinct discontinuous relationship between task difficulty and the optimal allocation of effort; people should refrain from engaging in very difficult tasks, but also invest more effort when success is marginally within reach. Despite its simplicity, the setting Kukla described is of broad interest, because it exposes the challenging nature of a common class of effort allocation problems encountered in daily life. Similar discontinuities between the parameters that characterize the decision-making problem (i.e., rewards, abilities, constraints) and the normatively prescribed or observed behavior (i.e., time allocation) have been uncovered in various domains, such as allocating time among learning tasks (Son & Sethi, 2006), effort in writing prestigious grants (Bol, de Vaan, & van de Rijt, 2018), and allocating attention in perceptual decisions (Morvan & Maloney, 2012), potentially hinting at a common underlying problem structure.

The notion that optimal solutions for even simple problems may harbor drastic discontinuities as a function of the task parameters contrasts with the main time allocation paradigms in economics and psychology, which assumes that different activities, often defined as effort and leisure, complement each other (Becker, 1965; Borjas & Van Ours, 2000; Kurzban, Duckworth, Kable, & Myers, 2013; Varian, 1978). Such resource allocation problems (called “convex” in the language of mathematics) are easy to solve using standard optimization techniques (Boyd & Vandenberghe, 2004), and the changes in the optimal solutions are smooth as we gradually shift the parameters of the problem. In a similar vein, optimal foraging theory (the most prominent time allocation paradigm in behavioral biology) is also a convex optimization problem where the foraging time allocated across different resource patches depends on immediate rewards and the rate at which these rewards diminish (Charnov, 1976). Although these modeling frameworks have been productively used to describe behavior and prescribe how to allocate resources across a wide range of domains, ranging from mental effort allocation (Kool & Botvinick, 2014) to cognitive search (Payne, Duggan, & Neth, 2007; Pirolli & Card, 1999), they cannot account for the behavioral dynamics produced by success–failure reward structures. In many real-life problems, returns directly depend on how additional efforts translate into improved chances of success in achieving a make-or-break goal.

Make-or-break tasks also provide a new and valuable perspective for studying risk-taking behavior in dynamic environments. Alternative theories of risk-taking have been empirically tested using experiments asking people to make consequential or hypothetical choices among a number of discrete risky options (Edwards, 1954b; Kahneman & Tversky, 1979; O’Donoghue & Somerville, 2018). Over the years, the models explaining empirical irregularities became increasingly more complex, adding further parameters and specifications to account for the largest number of choices. In real life, however, people take risks dynamically, and the decision to keep on investing in a risky challenge can be intimidating. Can risk-taking theories from the static domain be easily transferred to risky dynamic goal-pursuit problems? Accordingly, how could one implement an artificial

agent to be risk averse in the dynamic decision-making setting (Feldman & Sproull, 1977)? In challenging real-world environments, where calculating the optimal strategy is computationally excessive, existing expected and non-expected utility theories need to be cognitively simplified to become psychologically plausible (e.g., Ebert & Strack, 2015). Here, we explore one promising avenue. Risk-averse behavior, and other distinctive behavioral patterns (e.g., temporal discounting and perseverance), can be produced by leveraging boundedly rational principles, such as myopia or simple heuristics.

1.2. Contributions and scope

The first contribution of this article is to develop a formal framework that accounts for the role of uncertainty and information in resource allocation problems involving a make-or-break activity and a safe alternative. Our framework draws inspiration from Lewin's early work on aspiration levels (e.g., Lewin et al., 1944) and Kukla's (1972) static model of effort allocation, but with a radical departure in how uncertainty accumulates with larger investments of effort or time. We show that manipulating critical factors in the task, such as the rewards or constraints, can produce abrupt and discontinuous changes in the normative allocation strategy, with uncertainty playing a crucial role in determining the optimal solution. We describe two variants of the problem. In the one-shot allocation problem, the decision maker receives feedback about performance only at the very end of the task after all available time has been allocated. In this formulation of the problem, performance is unobservable during the allocation phase, as is the case when studying for a pass/fail exam or preparing a grant application. In the dynamic allocation problem, the decision maker receives immediate feedback on their current output and—like a PhD candidate toiling to fulfill graduation requirements before funding runs out—can dynamically adapt their allocation strategies at any point in time, either continuing to pursue the make-or-break goal or dropping out in favor of the safe alternative.

The second contribution of this article is to introduce the optimal solution for the dynamic version of the problem. In line with previous theoretical work on dynamic decision-making across disciplines in the behavioral and social sciences (Dixit & Pindyck, 1994; Houston & McNamara, 1999; Malhotra, Leslie, Ludwig, & Bogacz, 2018; McCardle, Tsetlin, & Winkler, 2016) we show that the solution can be expressed by optimal decision thresholds. We mathematically prove that optimally solving the dynamic allocation problem implies prioritizing investment in the make-or-break task over the safe rewards task (whenever investing in the make-or-break task is considered profitable) and switching unidirectionally to the safe alternative only when performance falls below the giving-up threshold or when success has been reached. This allows us to use well-known results from optimal stopping theory in stochastic processes (Shiryaev, 2007) as well as standard numerical methods to calculate the optimal giving-up strategy. Following previous work on optimal stopping, we discretize time and derive the solution by backwards induction. We show that acting optimally implies using a more tolerant giving-up threshold as uncertainty in the environment increases, echoing results from the dynamic investing literature in finance and economics (McDonald & Siegel, 1986). The optimal strategy

provides insights into the nature of the decision-making environment. Yet it is inaccessible to humans, due to its exorbitant computational costs (Simon, 1990). This raises the question of how to constrain the space of computational level theories in order to define cognitively plausible strategies that are available to human decision makers (Marr, 1982; Van Rooij, 2008).

The third contribution of this article is to formulate plausible boundedly rational strategies and to study their behavior and performance across decision-making environments. First, we define a *myopic giving-up* strategy, which is based on the optimal solution to the one-shot allocation problem. This myopic solution implies giving up earlier than the fully optimal strategy, because it does not consider the possibility of dynamically reassessing one's policy based on new information acquired during the task (i.e., direct feedback about performance). Myopic giving-up becomes more conservative relative to the optimal strategy as uncertainty increases, yielding a distinct pattern of risk-aversion. Second, we examine the *play-to-win* heuristic, a simple heuristic strategy that only decides whether or not to invest in the make-or-break task, and then either abandons the task entirely or stubbornly perseveres until success or failure occurs. When contrasted with optimal giving-up, this strategy produces risk-seeking behavior that is consistent with the sunk cost fallacy (Arkes & Blumer, 1985). Finally, we define the *control points* strategy, a generalization of the above two strategies, which considers giving-up at a fixed number of time points. The strategy allows for a more flexible trade-off of computational complexity and dynamic responsiveness. Holding all other factors constant, we find that an increase in uncertainty in the environment improves the relative performance of the play-to-win and control points strategies, which despite disregarding most information, almost always outperform the myopic giving-up strategy and can approximate the performance of the optimal solution.

The structure of this paper is as follows: We begin by developing the mathematical framework for the allocation problem and describing how different specifications of the decision environment can change the reward structure of the make-or-break task. We then describe the optimal solution to the one-shot allocation problem before moving on to the dynamic version of the problem. We present both optimal and myopic solutions to the dynamic problem, and we compare them to simpler heuristic strategies. In the Discussion, we draw connections with the dynamic decision-making literature from mathematical psychology and economics, with expected and non-expected utility theories, and with theories of effort allocation, motivation, and learning. These links highlight a rich set of opportunities for further research and contributions.

2. A formal framework

Let us consider an agent who can allocate time (or any another resource) t_m and t_s between two different reward-generating activities $\{m, s\}$, where m is an instance of a *make-or-break* task with a binary success-failure outcome, and s is an instance of a *safe* task, where rewards are a more predictable function of the allocated time. We assume that the total available time to be allocated T is fixed, with $T = t_m + t_s$. We introduce a

function describing how invested time maps onto *performance quality* (Eq. 1) and a function mapping performance quality onto *rewards* for each reward-generating activity (Eqs. 2 and 3).

Intuitively, the quality (or quantity) of performance $q(t)$ in any given task is a continuous function of the allocated time t and depends on the abilities of the agent, starting from $q(0) = 0$. Working on a task, however, is often accompanied by uncertainty about future performance. For instance, people may commit unintended and random errors (e.g., incorrect retrieval of information from memory during an examination), where these types of errors can be described using a stochastic function. It is reasonable to assume that the change in performance Δq over an interval of time Δt is a Gaussian random variable, independent of past performance, and that this random change Δq depends only on the magnitude of Δt (i.e., time homogeneity). The unique continuous function that satisfies all of the above assumptions is a diffusion process given by

$$q(t) = \lambda t + \sigma W_t. \quad (1)$$

The first term in Eq. 1 is a deterministic term and is proportional to time, where λ is the skill level of the agent. Larger values of λ lead to better performance for the same amount of time. The second term is a random term, where σ is the performance uncertainty defining the variance per unit of time, and W_t is the Wiener process. A Wiener process is the analog of a random walk for continuous time, where for each t , it yields a Gaussian random variable $W_t \sim \mathcal{N}(0, t)$, with a mean of 0 and variance equal to t . In the same way that the variance of a random walk increases with the number of steps, the variance of a Wiener process also increases with time t . As shown in Fig. 1, the expectation of performance quality (dotted line) is affected only by the skill level λ and the time t , whereas the underlying uncertainty (shaded area) is determined by the performance uncertainty σ and grows with larger investments of time t .

2.1. Rewards as a function of performance

Given a specific reward-generating activity (e.g., writing a grant application or working for an hourly wage), we can describe a function mapping performance onto rewards. Here, we consider two different types of functions, $g_m(x)$ for the make-or-break task and $g_s(x)$ for the safe task, where $x = q(t)$ is the performance function described in Eq. 1.

2.1.1. Make-or-break

In the *make-or-break* setting, the agent either receives a considerable reward B upon success, or nothing upon failure, depending on whether their performance $x = q(t)$ reaches a precise and known success threshold Δ :

$$g_m(x) = \begin{cases} B, & \text{if } x \geq \Delta, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

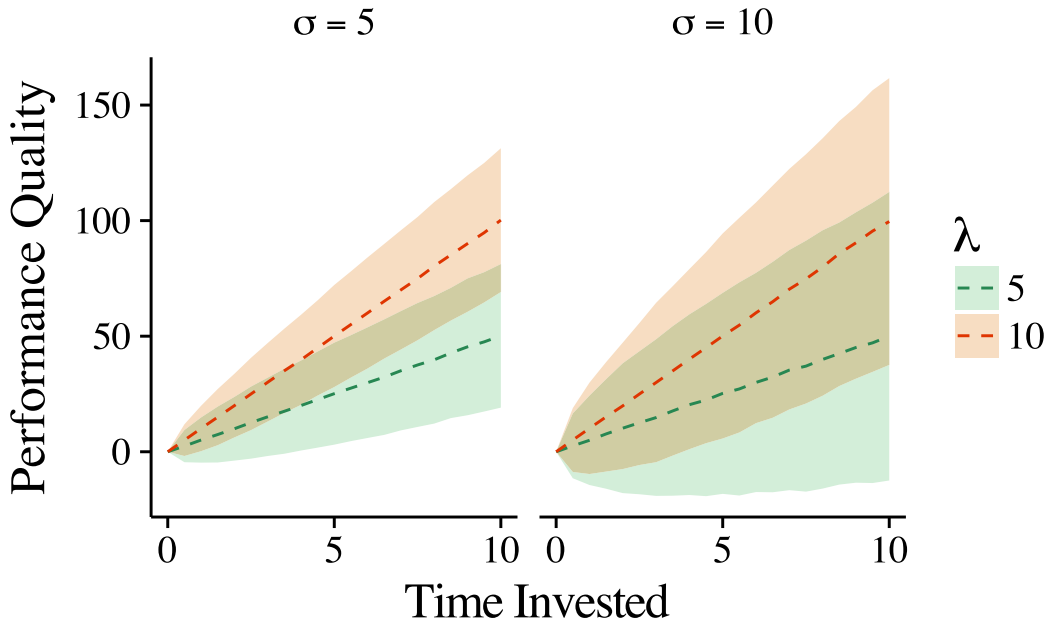


Fig. 1. Performance quality as a function of time invested, with total available time $T = 10$. The panels show different levels of performance uncertainty σ ; the colors represent different skill levels λ . Dotted lines indicate the expected performance quality and shaded areas show one standard deviation.

The performance threshold Δ captures the strict binary outcomes of success or failure and controls the difficulty of the make-or-break task. The reward B may consist, for example, in receiving a monetary bonus at work, winning an award, receiving a grant, or the joy of reaching an aspiration.² Note that it is not necessary for the reward to be a known value, but it could also be sampled from some known distribution, in which case B would denote the expected value of the reward.

2.1.2. Safe reward

In the *safe-reward* setting, rewards are a linear function of performance quality, with reward rate v corresponding to an hourly wage or a fixed rate of reward as a function of performance quality $x = q(t)$:

$$g_s(x) = v \cdot x. \quad (3)$$

2.2. Rewards as a function of resource allocation

We can now describe the stochastic processes of earning rewards as functions mapping time allocations t_m and t_s onto rewards r_m and r_s :

$$\begin{aligned}
r_m(t_m) &= g_m(q_m(t_m)) = g_m(\lambda_m t_m + \sigma_m W_{t_m}^m) \\
&= \begin{cases} B, & \text{if } \lambda_m t_m + \sigma_m W_{t_m}^m \geq \Delta \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{4}$$

$$r_s(t_s) = g_s(q_s(t_s)) = v \cdot q_s(t_s) = v \cdot (\lambda_s t_s + \sigma_s W_{t_s}^s) \tag{5}$$

The expectation of rewards for the make-or-break task can be written as

$$\begin{aligned}
\mathbb{E}[r_m(t_m)] &= B \cdot \mathbb{P}(\lambda_m t_m + \sigma_m W_{t_m}^m \geq \Delta) \\
&= B \cdot \mathbb{P}\left(\frac{1}{\sqrt{t_m}} W_{t_m}^m \geq \frac{\Delta - \lambda_m t_m}{\sigma_m \sqrt{t_m}}\right) \\
&= B \cdot \left(1 - \Phi\left(\frac{\Delta - \lambda_m t_m}{\sigma_m \sqrt{t_m}}\right)\right)
\end{aligned} \tag{6}$$

where $\mathbb{E}[\cdot]$ denotes expectation, $\mathbb{P}(\cdot)$ denotes probability, and Φ denotes the cumulative distribution function of a standard normal distribution. Regardless of the exact parameters of the problem, the above expression is always a sigmoid function of the allocated time t_m (see Supplementary Information Section 1). In contrast, the expectation of rewards for the safe alternative is

$$\mathbb{E}[r_s(t_s)] = v \cdot \lambda_s t_s. \tag{7}$$

Although uncertainty is a crucial component in calculating expected reward for the make-or-break task (Eq. 6), the same does not apply to the safe-reward task (Eq. 7), because the expected value of the Wiener process $W_{t_s}^s \sim \mathcal{N}(0, t_s)$ is always zero. Thus, the performance uncertainty σ_s does not influence the *expectation*, but only the variance of the reward. To give a better idea of these reward functions, Fig. 2 shows the influence of each parameter on expected reward for each activity type as a function of time allocation.

2.2.1. Sensitivity of expected rewards to environmental parameters

The expected rewards for the make-or-break task are a sigmoid function of time allocation (left column of Fig. 2), where the inflection point and the shape of the expected returns curve are jointly determined by the performance uncertainty σ , the skill level of the agent λ_m , the success threshold Δ , and the bonus B . First, note that the agent's skill level λ_m and the difficulty of the task (defined by the success threshold Δ) are directly related; increasing the former has a similar effect to decreasing the latter. Additionally, the performance uncertainty σ determines the shape of the expected returns in the make-or-break task where, in the degenerate case of $\sigma = 0$, expected reward becomes a step function of time allocation (middle left panel of Fig. 2). As uncertainty increases, the curvature of the sigmoid reward function becomes smoother. Finally, the bonus B in the

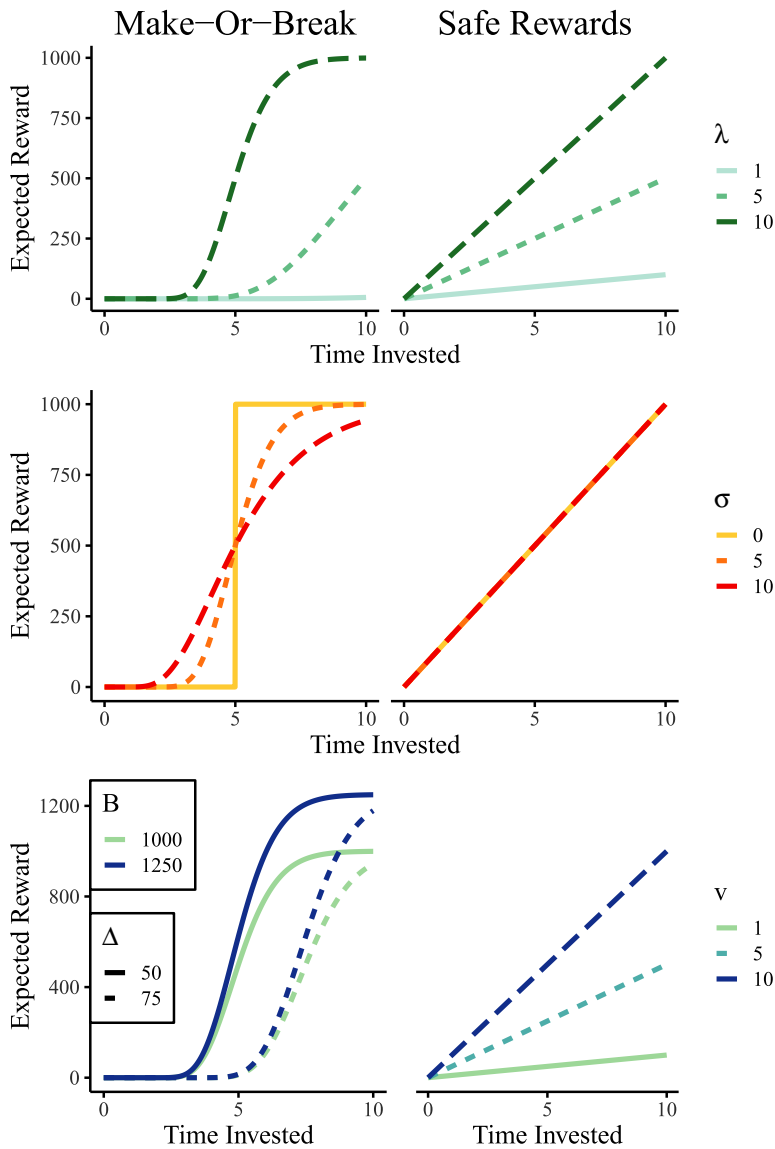


Fig. 2. Expected reward as a function of time invested in the make-or-break task (left column) and the safe-reward task (right column). The top row shows the influence of different skill levels (λ) and the middle row shows different levels of performance variance (σ). The bottom row shows the influence of the task-specific parameters: the bonus (B) and success threshold (Δ) bottom left, and the reward rate (v) bottom right. When not specified, we use the default parameter values of $T = 10$, $\lambda_m = \lambda_s = 10$, $\sigma_m = \sigma_s = 5$, $B = 1000$, $\Delta = 50$, and $v = 10$. These parameters are chosen such that the make-or-break reward is equal to the expected reward for the safe-reward task when all available time T is invested in the latter, that is, $B = v \cdot \lambda_s T$.

make-or-break task and the skill level λ_s and reward rate v in the safe-reward task control the relative payoffs between the two tasks. Multiplying or dividing both B and $v \cdot \lambda_s$ by the same amount does not alter the relative payoff in the tasks.

2.2.2. Marginal returns

Investing a small amount of time in the make-or-break task typically yields very small expected returns; the resulting performance is almost certain to fall below the success threshold. Initially, increasing time allocation in the make-or-break task corresponds to only small increases in expected reward, with returns being smaller when uncertainty is low. However, as one continues to invest more time in the make-or-break task, the rate of increase in expected reward accelerates and rises rapidly until it reaches an inflection point (at $t_m = 5$, middle left panel of Fig. 2). The rate at which expected reward increases around the inflection point is determined by performance uncertainty σ , with higher uncertainty creating a more gradual transition (smoother sigmoid curves; see middle left panel of Fig. 2). Eventually, as performance quality surpasses the threshold Δ , the marginal increases in expected reward become smaller and smaller, and the expected reward curve draws near the upper bound of B . Thus, after a certain expected performance level, additional effort has diminishing marginal returns.

3. Results

Having established the formal framework for the problem, we now present solutions to two versions of the problem and discuss their implications. First, we address the *one-shot* version of the problem, where the agent makes a single decision about how to allocate resources between the two tasks, without any feedback on performance or reward. Second, we address the *dynamic* version of the problem, where performance $q(t)$ is observable and switching between activities can occur at any point in time $t < T$ (i.e., after any amount of investment prior to exhausting all available time). We present a fully rational and a boundedly rational solution to the dynamic problem, both relying on the idea of a “giving-up” threshold; we then compare these solutions with simple heuristic strategies that achieve competitive performance at much lower computational costs.

3.1. The one-shot allocation problem

If an agent has a fixed amount of time T to allocate between the two tasks, and needs to commit to this allocation at the beginning of the task, how can they maximize the sum of expected rewards? This problem is identical to a case where the agent has to commit to a plan upfront, or to a scenario where there is no feedback on performance. Therefore, the optimization problem is reduced to a single decision about how to divide the total available time T between the two activities t_m and t_s , where the overall reward function can be written as:

$$\begin{aligned}
h(t_m, t_s) &= \mathbb{E}[r_m(t_m)] + \mathbb{E}[r_s(t_s)] \\
&= B \left(1 - \Phi \left(\frac{\Delta - \lambda_m t_m}{\sigma_m \sqrt{t_m}} \right) \right) + v \cdot \lambda_s t_s.
\end{aligned} \tag{8}$$

An agent should then optimize this function (Eq. 8) under resource constraints, which is written formally as follows:

$$\begin{aligned}
&\underset{t_m, t_s}{\text{maximize}} \ h(t_m, t_s), \\
&\text{subject to } T = t_m + t_s.
\end{aligned} \tag{9}$$

The above equations give rise to a simple, yet commonly encountered, non-convex optimization problem. Below, we describe and offer intuitions about how the optimal allocation policy depends on the parameters of the problem (see Fig. 3).

3.1.1. Playing safe or going all in

For some sets of parameters, one option completely dominates the other, and the decision maker should simply invest all available time resources T in just one of the tasks. In some cases, for instance, the rewards of the make-or-break task simply do not justify the risks, and the clear dominant option is to “play it safe” by investing all available time resources in the safe-reward task. This is the case when (a) there is only a slim chance of success in the make-or-break task (i.e., due to scarcity of total available time T , a very high threshold Δ , or low skill level λ_m ; see top left of Fig. 3, where $\lambda_m = 1$) or (b) when the expected returns from the safe-reward task are simply much higher (i.e., when $v \cdot \lambda_s T$ is much larger than B).

As we vary the parameters making the make-or-break task more attractive, it becomes rational for agents to allocate at least some—possibly all—of their time to the make-or-break task. There is a discontinuous switch in the optimal policy when a critical point in the parameter space is crossed as we increase the total available time, the bonus of the make-or-break task, or the skill level of the agent (see Section 2 in the Supporting Information). This is illustrated in Fig. 4, where the optimal policy switches from “playing it safe” ($t_s = T$) to “going all in” ($t_m = T$) at the critical value of λ_m satisfying the equation

$$B \cdot \mathbb{P}(\lambda_m T + \sigma_m W_T^m > \Delta) = v \cdot \lambda_s T. \tag{10}$$

Equation 10 corresponds to the point where the two extreme policies of “playing it safe” and “going all in” produce equal expected rewards. Transitions from one extreme to the other always occur when the make-or-break task is sufficiently rewarding (see Supporting Information for proof), and we vary the total available time resources T . In the Supporting Information, we specify the exact conditions under which all-or-nothing transitions should occur as we vary the skill level or the reward of the make-or-break task.

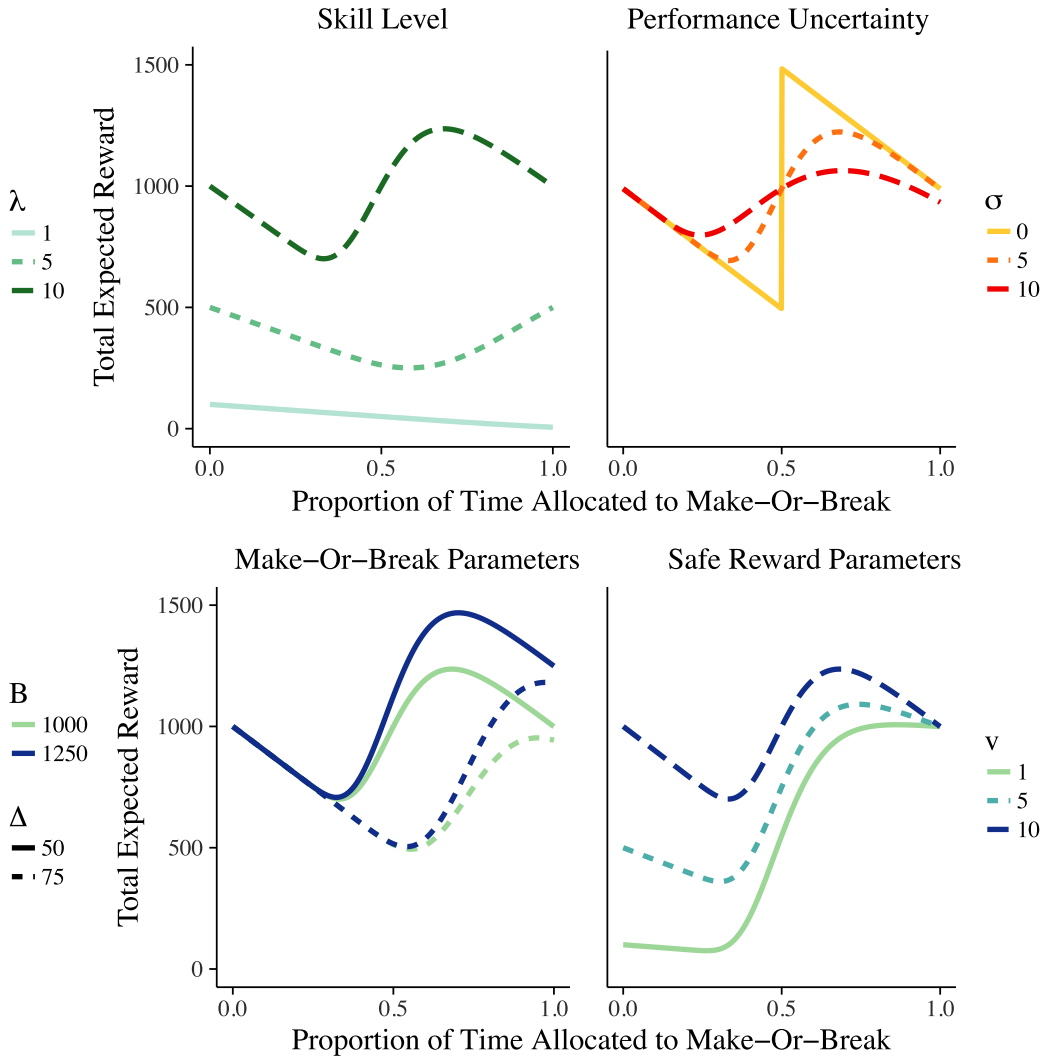


Fig. 3. Total expected reward as a function of the proportion of time allocated to the make-or-break task, where remaining time is allocated to the safe-reward alternative. The top left panel shows how the shape of the curve varies for different skill levels (we set $\lambda_m = \lambda_s = \lambda$). The top right panel shows the effect of uncertainty (in the make-or-break task, $\sigma = \sigma_m$) on the curve shape. The bottom panels show how the returns curve changes as we vary the bonus or the threshold (left) or the reward rate in the safe task (right). When not specified, we use the following default parameter values: $T = 10$, $\lambda_m = \lambda_s = 10$, $\sigma = 5$, $B = 1,000$, $\Delta = 50$, $v = 10$.

3.1.2. Slacking off for the highly skilled

In environments where success in the make-or-break task is highly likely without requiring the investment of all available time, the optimal policy entails allocation of resources to both activities. As expected performance in the make-or-break task increases

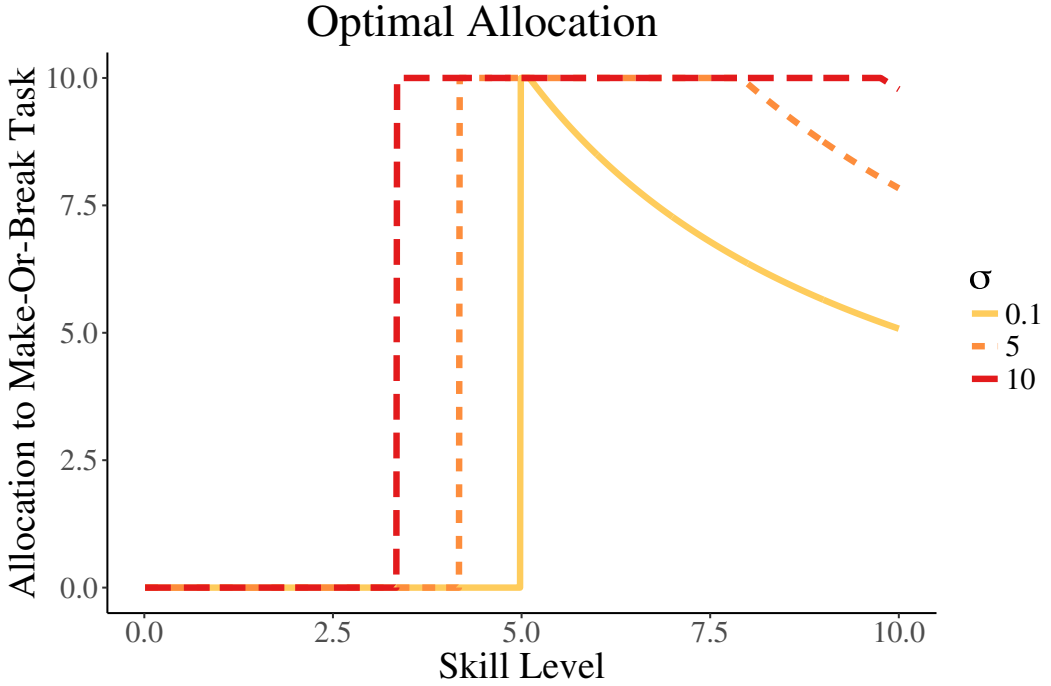


Fig. 4. Optimal allocation as a function of skill level and under different levels of performance uncertainty (line type). Unless otherwise specified, we use the following default parameter values: $T = 10$, $B = 1000$, $\Delta = 50$, $v = 10$. To illustrate the crucial role of uncertainty, we set $\lambda_s = 3$ and vary the value of $\lambda_m \in [0, 10]$. For small performance uncertainty ($\sigma = 0.1$), it is profitable to switch from allocating everything to the safe-reward task to allocating everything to the make-or-break task when an agent has a 50% chance of success. At higher uncertainty levels, this critical point occurs at a lower skill level (compare the optimal allocation policy for $\sigma = 0.1$ with $\sigma = 5$ and $\sigma = 10$). On the other hand, people with higher skill levels should allocate time to the make-or-break task until the point that marginal returns from increasing the chances of success in the task are equal to the linear opportunity cost of the safe-reward task (see Eq. 11). At higher uncertainty levels, this critical point occurs at higher skill levels.

beyond the critical point described in Eq. 10, the decision maker can potentially err on the side of investing too much time in the make-or-break task. This error comes with the opportunity cost of losing potential rewards from investing surplus time in the safe-reward task. This is the case when the marginal expected gain of investing more time in the make-or-break task is less than the gain of investing more time in the safe-reward task, which is as follows:

$$[E[r_m(t_m)]]'_{t_m=T} < v \cdot \lambda_s, \quad (11)$$

where the derivative is taken with respect to the allocation t_m . In Fig. 4, this corresponds to the point at which the optimal allocation of time to the make-or-break task begins to decrease, as the likelihood of success increases as a function of skill level (λ_m).

Everything else being equal, a lower threshold and more total available resources make it more likely that an agent should allocate time to both activities. People with higher skill levels can quickly secure a higher probability of success and invest residual time resources in the safe-reward task. Whereas Kukla (1972) suggested a discontinuous relationship between the difficulty of a task and the effort people exert, here we have generalized it to any level of uncertainty and clarified the relations between the different parameters of the problem (see Fig. 4, where skill and difficulty are inverse concepts).

3.1.3. *The motivating effect of uncertainty*

Everything else being equal, higher performance uncertainty implies that it is optimal to allocate all time to the make-or-break task for a larger subset of parameters in the parameter space. People with lower skill should dedicate themselves completely to the make-or-break goal in hope of eventual success. It also makes sense for people with higher skill to put all their effort in the make-or-break task, in order to minimize the risk of failing (see also Fig. 1 in the Supporting Information). Higher uncertainty leads to an increase in the expected rewards of the make-or-break task for small amounts of time allocation, and it also implies that the deceleration of marginal returns occurs more slowly (see the middle left panel of Fig. 2 and the upper right panel of Fig. 3). As a result, in more uncertain environments (i.e., larger σ), the critical point described by Eq. 10 occurs for lower levels of skill or total available time, while the critical point described by Eq. 11 occurs for larger values. Fig. 4 demonstrates this effect when comparing a low uncertainty environment ($\sigma = 0.1$) with intermediate ($\sigma = 5$) or high uncertainty environments ($\sigma = 10$). The motivating effect of uncertainty is particularly pronounced when comparatively large rewards are at stake in the make-or-break task (i.e., a high $B / (v \cdot \lambda_s T)$ ratio) and gradually attenuates as the relative rewards from the safe-reward task increase.

3.1.4. *Discussion*

In everyday experience, people often need to decide how to allocate their time or effort between challenging make-or-break goals and safe alternatives. We have shown that this type of resource allocation problem is non-convex in nature, leading to striking discontinuities in how people should allocate their time as the parameters of the problem vary. Small differences in people's skill levels, resources, or the rewards available may change the optimal policy from allocating all of one's time to the safe-reward task to investing everything in the make-or-break task. Crucially, we have shown how the degree of uncertainty in the environment moderates the optimal allocation policy and determines when these discontinuities occur.

3.2. *The dynamic allocation problem*

So far we have assumed that the agent makes a single allocation decision, dividing the available time between two activities without any feedback on performance. In many problems, however, people dynamically obtain information about their performance,

which they can use to reassess their prospects of success and alter their behavior on the fly. How should agents dynamically invest their time when they can directly observe their performance and switch dynamically between tasks? In the analyses that follow, we assume that the agent has full knowledge of all the critical parameters of the problem and his current performance but is uncertain only about his future performance.

3.2.1. Optimal allocation policy

3.2.1.1. Order of tasks and switching between tasks: Intuitively, people above a certain skill level in the make-or-break task should start by investing time toward it; they can switch to the safe-reward task in case of an early success or if success seems unattainable. Because future rewards in the safe reward task do not depend on past performance (i.e., there is no useful feedback from the safe task), the dominant strategy is to begin with the make-or-break task and to switch unidirectionally to the safe-reward task (see Supporting Information Section 3 for a proof). This leads to a significant simplification of our problem: An optimal solution to the allocation problem involves at most one switch between tasks and only from the make-or-break to the safe-reward task. The question then is when to make this single switch.

3.2.1.2. Giving-up threshold: At any given point in the make-or-break task, the agent has to decide whether to switch immediately to the safe-reward task or to further pursue the make-or-break goal, based solely on the current performance. At very poor performance levels, success is highly unlikely; therefore, a rational agent should switch to the safe-reward task. On the other hand, if the agent is relatively close to the success threshold (and assuming the reward is large enough), it would be sensible to continue in pursuit of eventual success. For a specific intermediate level of performance, the two strategies will yield the same expected reward, such that the agent can be indifferent about which strategy to follow. We call this indifference point the *giving-up threshold* and at time t we denote it by c_t . Intuitively, the agent should abandon the make-or-break task in favor of the safe-reward task if, and only if, performance is equal to or falls below this value after time t . So long as the agent starts with a performance above the giving-up threshold (i.e., $q_m(0) > c_0$ such that it is rational to invest in the make-or-break task at all), task switching should occur when performance $q_m(t)$ falls below the giving-up threshold c_t or upon reaching the success threshold Δ . More precisely, the point at which an agent should give up is given by

$$\tau = \min\{t : q_m(t) \leq c_t \text{ or } q_m(t) \geq \Delta\}. \quad (12)$$

In the stochastic processes literature, this type of problem is considered a problem of *optimal stopping*. The existence of a giving-up threshold that leads to an optimal policy through Eq. 12 is guaranteed by Theorem 2.2 in Peskir and Shiryaev (2006), with several methods for deriving this optimal policy given in Chapter 8 of the same book.³ To the best of our knowledge, there is no analytic solution for finding the optimal policy and all approaches that can be used to find it are computationally expensive. Here, we follow the

most widely used approach for deriving the optimal policy, which involves discretizing time and using backwards induction (Dixit & Pindyck, 1994; Kushner & Dupuis, 2013). In this section, we use the expected value theory framework in order to convey our results with the greatest clarity. Yet our analyses can also be generalized to the framework of expected utility theory, and it is possible to account for risk-aversion, risk-seeking, or inter-temporal discounting (see Supporting Information Sections 5 and 6).

First, we discretize time by dividing T into n equally sized intervals. This turns the problem into a discrete optimal control problem whose solution involves solving, using backwards induction, the resulting Bellman Equation

$$R_k(x) = \max \left\{ \int_{-\infty}^{\infty} R_{k+1}(y) \cdot \phi_x(y) dy, g_m(x) + v \cdot \lambda_s \cdot \frac{(n-k)T}{n} \right\}, \quad (13)$$

for $k = 0, \dots, n-1$. $R_k(x)$ is the expected total reward from time k to n , if at time k the performance in the make-or-break task is x , and ϕ_x is a suitable normal distribution. The initial conditions are given by $R_n(x) = g_m(x)$ (see Supporting Information Section 4 for details). Thus, the giving-up threshold c_k at time k is the unique solution of the equation:

$$\int_{-\infty}^{\infty} R_{k+1}(y) \cdot \phi_{c_k}(y) dy = v \cdot \lambda_s \cdot \frac{(n-k)T}{n}. \quad (14)$$

The optimal policy is to switch to the safe reward task at the first step k such that $q_m(t_k) \leq c_k$, due to falling below the threshold, or $q_m(t_k) \geq \Delta$, due to success. Deriving the optimal threshold requires starting from the end of the task and reasoning backwards. Thus, calculating the giving-up threshold at t_0 requires knowledge of the threshold values at all subsequent steps. Finding the optimal solution is computationally demanding not only for humans but also for machines, as we explain in the following paragraph.

3.2.1.3. The computational burden of calculating the optimal threshold: Herbert Simon, one of the founding fathers of artificial intelligence, argued that theories of rationality that do not take into account the computational costs of problem solving are incomplete or even misleading (Simon, 1978). We build on his conception of rationality and apply computational complexity theory (Papadimitriou, 2003; Van Rooij, 2008) to express the computational cost of the optimal stopping strategy. As shown in Eq. 13, for each step k and each performance value x , we have to calculate an integral over all performance values y at the next step. Assuming that the integral is computed numerically and approximated by a sum, both its computational cost and the accuracy of the result depend on the number of summands we use. If we use m terms in the sum, then the computational complexity of one integration will be $\mathcal{O}(m)$. This will give us $R_k(x)$ for a specific k and x . But in order to later find $R_{k-1}(x')$, for any x' , we need to have $R_k(x)$ for m different values of x . Thus, for each k , we need $\mathcal{O}(m^2)$ computations. This gives us $R_k(x)$ for a specific k and for m different values of x . The accuracy of the approximation also depends on the

number of steps we use in the time axis, that is, the number of values of k . Thus, if we divide time into n steps, we need in total $\mathcal{O}(n \cdot m^2)$ computations for the solution described above, which results in an algorithm of polynomial complexity. Although polynomial-time algorithms are considered tractable in computer science and artificial intelligence, as opposed to non-polynomial-time (for example, exponential-time), in practice they can still be prohibitively expensive to compute,⁴ even for exponents as low as 2, like $\mathcal{O}(m^2)$. Even for relatively small values for n and m (e.g., $n = m = 100$), the optimal policy would require in the order of a million computations ($n \cdot m^2 = 10^6$). Regardless of the exact cost measure used to penalize computation, the demands of the optimal strategy are excessive (for different operationalizations of computational costs see Fechner, Schooler, & Pachur, 2018; Johnson & Payne, 1985; Lieder & Griffiths, 2017). The computational burden would be even larger for an optimally behaving risk-averse or risk-seeking agent, since additional parameters would need to be computed (see Supporting Information Section 6). An important principle in cognitive science is to constrain the space of computational level theories to plausible strategies at the algorithmic level (Marr, 1982). Thus, it becomes clear that alternative strategies with lower computational complexity would provide valuable insights into how human decision makers grapple with these problems or even hint to how to program artificial agents coping with similar decision problems.

3.2.2. *Boundedly rational alternatives*

Dynamic decision-making problems are both widely prevalent (Brehmer, 1992; Friedman, Isaac, James, & Sunder, 2014; Gonzalez, Fakhari, & Busemeyer, 2017; March, 1996) and notoriously challenging from a computational perspective. While optimal solutions often depend on backwards induction, several studies have shown that even for shallow planning depth, humans do not use backwards induction (e.g., Hotaling & Busemeyer, 2012; Huys et al., 2015; Zhang & Yu, 2013). The same holds in game-theoretic contexts (Johnson, Camerer, Sen, & Rymon, 2002; McKelvey & Palfrey, 1992). Instead of using computationally complex solutions, people tend to rely on simple and computationally inexpensive algorithms. While heuristics can lead to consistent deviations from optimality in some environments (Lieder & Griffiths, 2017; Tversky & Kahneman, 1974), they can produce competitive performance in other environments (Analytis, Kothiyal, & Katsikopoulos, 2014; Gaissmaier & Schooler, 2008; Gigerenzer & Todd, 1999). In the following sections, we examine several boundedly rational strategies that circumvent the computational costs of full rationality, yet perform competitively in some environments.

3.2.2.1. *The myopic giving-up strategy:* An alternative to backward induction is to assume that the decision maker acts myopically and decides whether or not to pursue the make-or-break task further as if this decision were the last that they could make. Myopic strategies accurately describe human behavior in a wide array of settings, ranging from dynamic investment decisions and sequential hypothesis testing to sequential search and multi-armed bandit tasks (see also Busemeyer & Rapoport, 1988; Gabaix, Laibson,

Moloché, & Weinberg, 2006; Stojic, Analytis, & Speekenbrink, 2015; Thaler, Tversky, Kahneman, & Schwartz, 1997; Wu, Schulz, Speekenbrink, Nelson, & Meder, 2018; Zhang & Yu, 2013). In our case, as more time is allocated to the make-or-break task, some of the associated uncertainty becomes replaced by an actual outcome $y = q_m(t')$ experienced up until time $t' \in [0, T]$. The agent can use this information to revise their allocation policy, making a myopic decision at any point during the allocation problem. Because future performance is independent of past performance (by the properties of the Wiener process), the problem of finding the optimal allocation is the same as before, but with total time $T_r = T - t'$ and threshold $\Delta - y$. The expected reward for the make-or-break task, if additional time t_m were invested in it, would be

$$\begin{aligned} \mathbb{E}[r_m(t_m)] &= B \cdot \mathbb{P}(\lambda_m t_m + \sigma_m W_{t_m}^m \geq \Delta - y) \\ &= B \cdot \left(1 - \Phi\left(\frac{\Delta - y - \lambda_m t_m}{\sigma_m \sqrt{t_m}}\right) \right). \end{aligned} \quad (15)$$

The expected returns for the safe-reward option would still be $E[r_s(t_s)] = v \cdot \lambda_s t_s$. A myopic agent seeks to optimize the sum of these rewards, under the constraint $T_r = t_m + t_s$, and will continue to invest in the task for as long as it would myopically be profitable to do so. A giving-up threshold in terms of performance can be computed (Fig. 5, dashed line) as the point below which the myopic policy prescribes investing no further time in the make-or-break task, in other words when optimally $t_m = 0$. When performance drops below this threshold, the myopic agent will switch to the safe-reward task.

Finding the t_m that maximizes the total expected reward is an easy optimization problem, which can be solved by finding the roots of the derivative of Eq. 15. The computational complexity of this, using Newton's method, for example, is a constant multiple of the complexity of computing elementary functions (Brent, 1976). Hence, at any point, an agent can decide whether to continue with the make-or-break task by solving a computationally inexpensive problem. In total, $O(n)$ calculations will be needed, but distributed equally over n intervals. This starkly contrasts with the optimal strategy, where the entire computational cost has to be paid upfront. The myopic strategy produces highly risk-averse behavior. In Fig. 6 we also compare the myopic strategy with an optimally behaving risk-averse agent whose preferences are described by a power function. As uncertainty in the environment increases, the myopic strategy aligns with progressively higher levels of risk-aversion.

3.2.2.2. The play-to-win strategy: Another way to cope with the computational complexity of dynamic decision-making is to employ heuristic strategies (Gigerenzer & Gaissmaier, 2011; Tversky & Kahneman, 1974). We first consider the play-to-win strategy, a heuristic strategy that stubbornly pursues success in the make-or-break task, switching to the safe-reward task only when the make-or-break threshold is reached. Thus, the strategy only needs to decide whether to start pursuing the make-or-break goal at all, eliminating

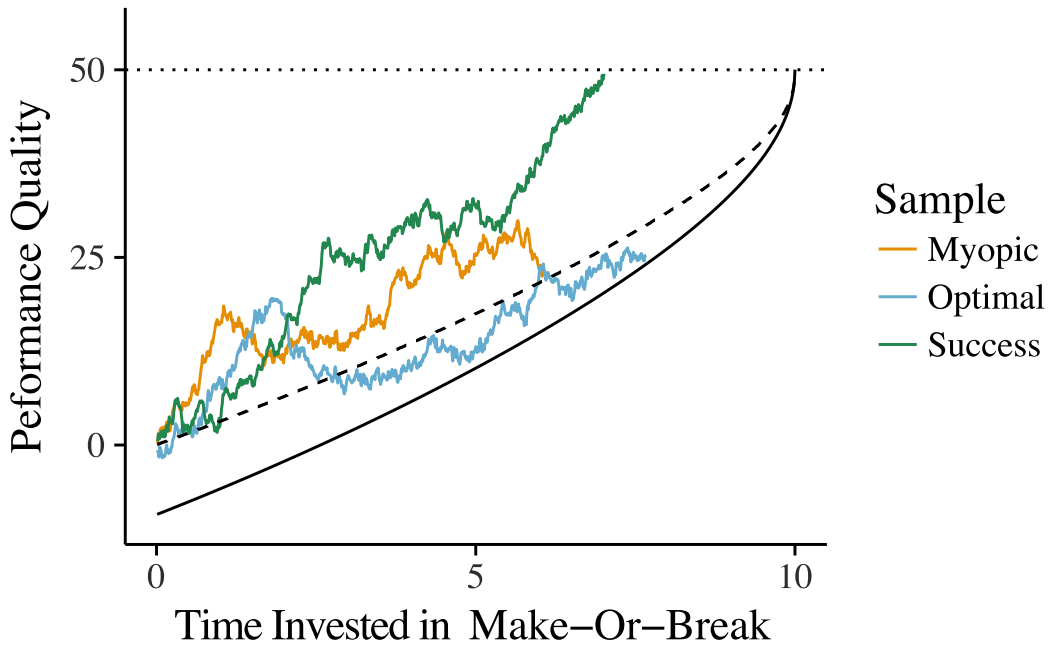


Fig. 5. Visualization of the *myopic* (dashed line) and *optimal* (solid line) giving-up thresholds for three samples drawn from the stochastic performance process $q(t)$. The green sample reaches the success threshold (dotted line) and is free to invest the remaining time in the safe-reward task. The orange sample has hit the myopic giving-up threshold (dashed line) and therefore gives up on the make-or-break task. However, an optimal decision maker has a more tolerant giving-up threshold (solid line); thus, the blue sample and the green sample continue even after hitting the myopic threshold, resulting in either a later giving-up point (blue) or eventually reaching the success threshold (green). Any time resources remaining after the agent reaches either a giving-up threshold or the success threshold are allocated to the safe-reward task. We use the following default parameter values: $T = 10$, $\sigma_m = \sigma_s = 5$, $B = 1000$, $\Delta = 50$, $\nu = 10$, but use $\lambda_m = \lambda_s = 5$ for illustrative purposes.

the computational costs almost entirely. Although the strategy is suboptimal, it has an intuitive appeal. Most people are familiar with real-world examples of individuals who, once they had taken the first step, would not give up on a goal, irrespective of their chances of success. The crucial event calibrating the expected returns on this strategy is the first time point at which the stochastic process crosses the success threshold. This event is commonly referred to as the first passage time of the stochastic process. First passage time problems for a Wiener process with a single threshold have been studied extensively across scientific disciplines (see Redner, 2001) and the results can be readily transferred to our setting. To calculate the expected total reward for this strategy, we need to bear in mind that there are two cases: either (a) the agent pursues the make-or-break task, but never reaches the reward threshold Δ and receives zero reward, or (b) the agent reaches the reward threshold after some time $\tau \leq T$ and receives the bonus B in addition

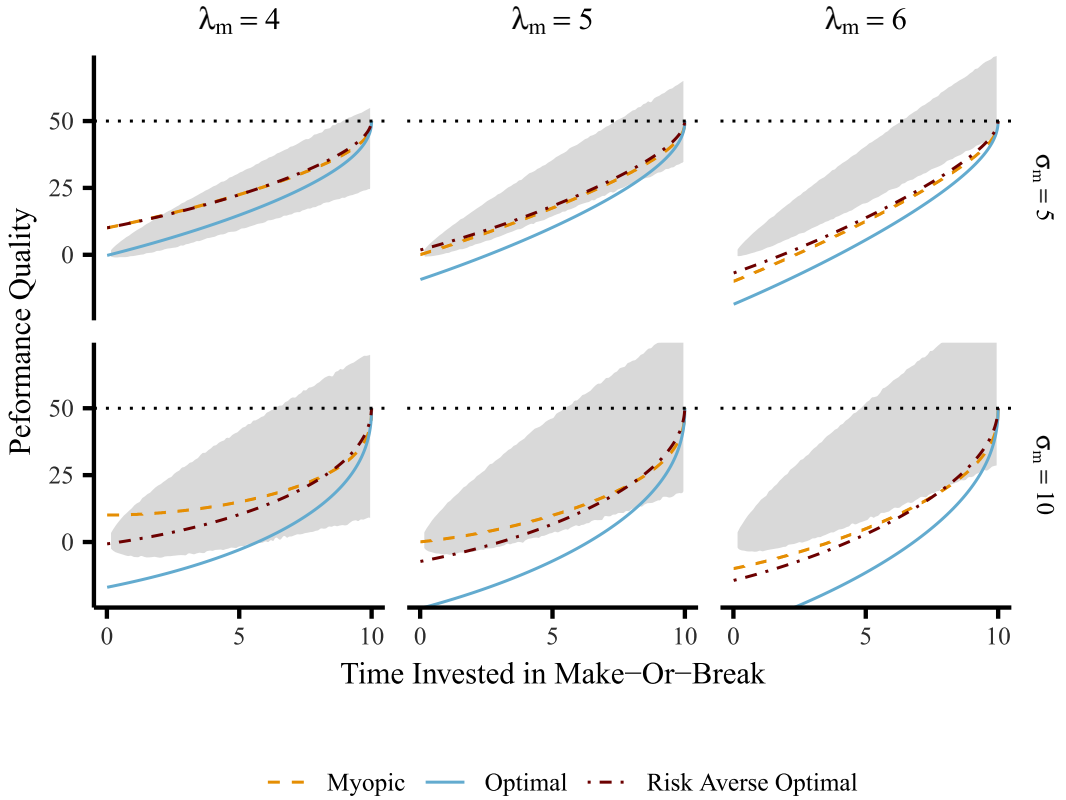


Fig. 6. Myopic, optimal, and risk-averse optimal ($\rho = 0.75$) giving-up thresholds, as a function of performance variance (σ_m) and ability (λ_m) in the make-or-break task. Other parameters are defaulted to $T = 10$, $\sigma_s = 5$, $B = 1000$, $\Delta = 50$, $v = 10$, and $\lambda_s = 5$. The shaded area shows one standard deviation of performance quality. As σ_m increases, the differences between the myopic and optimal thresholds increase. Note that the myopic giving-up strategy becomes progressively more risk-averse under higher uncertainty (i.e., more risk-averse than the $\rho = 0.75$ optimal risk-averse curve when $\sigma_m = 10$). See Supporting Information Section 6 for how to incorporate risk preferences in our framework.

to rewards from the safe-reward task, to which any surplus resources $T - \tau$ were allocated. The time τ can be defined mathematically as follows:

$$\tau = \min\{t : q_m(t) \geq \Delta\} = \min\left\{t : W_t^m + \frac{\lambda_m}{\sigma_m} \cdot t \geq \frac{\Delta}{\sigma_m}\right\}. \quad (16)$$

In the stochastic processes literature, the above is referred to as a hitting time or first passage time, at level $\frac{\Delta}{\sigma_m}$, of a Wiener process with drift $\frac{\lambda_m}{\sigma_m}$. We provide the exact probability density function of τ (known as an inverse Gaussian distribution) and an analytical expression for the expected reward for this strategy in Supporting Information Section 7.

Intuitively, the play-to-win heuristic can be ruinous for agents who start when their chances of succeeding in the make-or-break task are negligible. To perform competitively

across environments, it has to be combined with a good starting rule. What is a good starting rule for the play-to-win strategy? For somebody in possession of a calculator and the analytical formula for the expected returns on the strategy (Supporting Information Section 7), it would be trivial to compute whether the expected returns are higher for the make-or-break task than for playing it safe. Even if suboptimal, this rule ensures that the agent will reap—whenever possible—the higher expected returns procured from stubbornly pursuing the make-or-break goal. However, it is unlikely that laypeople will be able to use the analytical formula. Alternatively, an individual could employ a myopic calculation (using Eq. 15 only at $t = 0$) to gauge whether to start pursuing the make-or-break goal in the first place, and then never give up. Such a starting rule is more cognitively plausible and would protect people from starting in contexts where success is unlikely (e.g., where the decision maker has a low skill level and the uncertainty in the environment is relatively low), eliminating the costs of continuously monitoring their progress and calculating a giving-up threshold. Note that in order to illustrate both the adaptive and pathological use of the play-to-win heuristic, we simulate the strategy in Fig. 7 without a starting rule, but we also discuss the psychologically plausible myopic starting rule.

3.2.2.3. Control points strategy: Another plausible heuristic strategy is for the agent to consider the possibility of giving up at only a fixed number of time points (e.g., at the end of the academic or financial year). According to this strategy, there are two cases in which the agent may switch to the safe-reward task: if she passes the reward threshold Δ or if at one of the “control points” her performance drops below the giving-up threshold c_t . More precisely, the time τ that the agent will switch to the safe-reward task is

$$\tau = \min\{t : q_m(t) \geq \Delta \text{ or } (t = t_k, t_k \text{ is a control point and } q_m(t_k) \leq c_{t_k})\}. \quad (17)$$

To some degree, educational institutions and businesses already implement this strategy. They often require members to pass interim performance thresholds for progressing to the next level, rather than supervising them continuously. The strategy can be also seen as a self-control strategy (e.g., Ainslie, 1992; Metcalfe & Mischel, 1999), since people only have to consider giving up at a few discrete points in time rather than face the constant temptation to quit (as is the case for the myopic threshold).

The play-to-win and myopic giving-up strategies are special cases of the control points strategy. The control points strategy considers giving-up at N points in time, while in the case of the play-to-win strategy, the agent considers giving-up only at $N = 1$ point situated at the beginning of the task $t = 0$. On the other extreme, if we let $N \rightarrow \infty$, we recover the myopic strategy. Although there are many possible ways one can choose the points of control and the threshold, we will consider only a simple one, where the control points are placed at equal intervals and the intermittent thresholds are calculated using a myopic calculation. For instance, if the total available time is 10 units and there are 4 control points, then these will be at $t \in \{2, 4, 6, 8\}$ and the threshold at these points will be set according to Eq. 17.

3.2.3. Strategy comparison

For the play-to-win heuristic, the expected returns and the distribution of the time of success can be derived analytically. In contrast, for threshold-based strategies and the control points heuristic, the expected returns can be derived only numerically and the distribution of success and dropout times can only be determined using simulations. To obtain additional insights into how the strategies compare to one another, we therefore simulated the behavior of decision makers in the make-or-break task at three skill levels ($\lambda_m \in \{4, 5, 6\}$) and two levels of uncertainty ($\sigma_m \in \{5, 10\}$). These conditions are sufficient to capture the crucial factors influencing the performance of the strategies as well as the interaction between them. Fig. 7 compares the simulated behavior of the various strategies in order of complexity (decreasing from left to right), showing the distribution of attained rewards from which the timing distribution of successes and drop-outs can be derived. We implement the play-to-win heuristic without a starting rule (i.e., it always pursues the make-or-break task) for illustrative purposes, and implement the control points strategy with 4 control points at $t \in \{2, 4, 6, 8\}$, where intermittent thresholds are calculated myopically. Table 1 provides a short summary of each strategy.

3.2.3.1. Optimal versus myopic giving-up strategies: Looking first at the optimal and myopic giving-up strategies, the differences in performance shown in Fig. 7 can be explained by a divergence in their respective thresholds (Fig. 6). As uncertainty in the environment increases, so does the value of information that is received through feedback in the course of the task. Only the optimal giving-up strategy integrates the value of information that will be revealed in the future and thus prescribes persevering longer (before dropping out), whereas the myopic giving-up drops out more conservatively. At very low or very high skill levels, the optimal and myopic giving-up strategies lead to similar performance. At very low skill levels, both strategies play it safe; at very high skill levels, even the more conservative myopic giving-up threshold is rarely ever triggered.⁵ We now focus on discussing how skill and uncertainty influence the performance of the optimal or myopic giving-up strategies. Everything else being equal, the results will be more pronounced when a larger reward is at stake (for high $B / (v \cdot \lambda_s T)$ ratios), and the differences will attenuate when the relative rewards from the safe-reward task increase. Here we present results for $B = 2v \cdot \lambda_s T$. The analysis could be also conducted for inherently risk-averse or risk-seeking agents. The main difference then is that as we increase the risk-aversion in the utility function the relative performance of the myopic strategy as compared to the other strategies improves, making it more attractive for risk-averse decision makers.

The effect of skill level: The divergence between the myopic and optimal giving-up strategies is most pronounced at intermediate skill levels (i.e., $\sigma = 5$, $\lambda_m = 5$). Under these conditions, it is only marginally profitable (from a myopic perspective) to allocate any amount of time to the make-or-break task. The myopic threshold is rather conservative, leading to quick dropouts in most simulation runs. Agents following the myopic strategy succeed in the make-or-break task in very few cases (3.1%). In contrast, simulated agents following the optimal strategy persevere for longer and succeed in

Table 1
Summary of strategies

Strategy	Summary	Complexity
Optimal	Derived through backwards induction and implemented as a giving-up threshold calculated for each time point. If performance falls below the threshold, the optimal strategy dictates that the agent should give up and switch to the safe alternative.	$\mathcal{O}(n \cdot m^2)$
Myopic	Similar to the optimal strategy, but using only a single forward-looking myopic calculation to determine the giving-up threshold. This myopic calculation implies a more conservative threshold as uncertainty increases.	$\mathcal{O}(n)$
Control points	Rather than computing a giving-up threshold at all possible time points, the control points strategy only considers giving up at a fixed number of k time points using forward-looking calculations. In the limit of infinite control points, it becomes the myopic strategy.	$\mathcal{O}(k)$
Play to win	This is a fixed strategy that ignores feedback about performance, and stubbornly pursues the make-or-break task until eventual success or failure. This strategy can also be combined with a number of different starting rules that gauge whether the make-or-break task is worthwhile to pursue until an eventual success or failure.	$\mathcal{O}(0)$ without a starting rule and $\mathcal{O}(1)$ with a starting rule.

Note: n denotes the number of discrete time points, m are the discretized number of possible performance levels, and k is the number of control points considered, where $k < n$.

approximately half the trials, with dropouts more evenly distributed across time (see the upper middle panel of Fig. 7). At relatively low skill levels (e.g., $\sigma_m = 5$, $\lambda_m = 4$), an optimal decision maker would pursue the make-or-break task, whereas a myopic agent would play it safe. Even under these conditions, an unexpectedly good start could bring the success threshold within reach for an optimal decision maker. If performance falls below expectations, a decision maker can switch to the safe-reward alternative, thereby acquiring at least some rewards. Only a few of the simulated agents following the optimal strategy persevere until success (only 3.3% for $\sigma_m = 5$, $\lambda_m = 4$), with most simulated agents dropping out early (upper left panel of Fig. 7). Under this set of parameters, the optimal policy leads to a modest increase in terms of expected returns, where rare successes are largely balanced out by early dropouts. The difference between the two giving-up strategies is still notable—yet much smaller—at relatively high skill levels (see the upper right panel of Fig. 7).

The effect of uncertainty: The two threshold strategies diverge even more in terms of performance and observed behavior as uncertainty increases for two reasons. First, higher uncertainty implies a higher value of dynamically revealed information. The optimal strategy integrates this information and becomes more tolerant overall, investing in the make-or-break task at even lower skill levels. Second, higher uncertainty leads to larger fluctuations in performance, causing myopic agents to prematurely drop out more often,

even at higher skill levels. Both these factors make the optimal threshold much more perseverant at high levels of uncertainty when compared to the conservative myopic threshold (contrast the upper and lower panels of Fig. 7).

3.2.3.2. When and why you should never give up: For moderate or high skill levels in the make-or-break task, the play-to-win heuristic does fairly well relative to the optimal strategy in terms of average rewards and can even outperform the myopic strategy. As uncertainty in the environment increases, the play-to-win heuristic approximates the expected rewards of the optimal strategy (see Fig. 7, middle and right columns for $\sigma = 10$ and $\lambda_m = 5, 6$). How is this possible, given the stark difference in computational costs? In environments with high uncertainty, the optimal strategy is already more tolerant in order to accommodate the increased value of information. Yet even in cases where performance falls below the optimal giving-up threshold, it is not unlikely that performance will rebound and ultimately reach the success threshold. In environments with extreme uncertainty, the play-to-win heuristic approximates the expected rewards of the optimal giving-up strategy, rendering the computational costs for deriving the optimal threshold an unnecessary burden. Recall that for illustrative purposes, Fig. 7 shows the play-to-win strategy without a starting rule, such that it always pursues the challenging goal. This results in pathological cases, where play-to-win pursues the make-or-break task even when the prospects of success seem impossibly bleak at the outset (see upper left panel of Fig. 7 when $\sigma_m = 5$ and $\lambda_m = 4$, where the play-to-win strategy would succeed in only about 27% of the cases, and gain no rewards in other cases). This outcome can easily be prevented by adding a starting rule using a myopic calculation to assess whether it is worthwhile at all to invest in the make-or-break task, in which case effort is continually allocated until success or failure ensues. With a myopic starting rule, the play-to-win strategy outperforms the myopic threshold strategy in all examined environments where the make-or-break task is myopically perceived as attractive, despite disregarding feedback about performance.

The benefits of giving up full control. The control points strategy largely outperforms the myopic strategy despite being computationally cheaper. It also slightly outperforms play-to-win, further closing the gap to the optimal strategy, but at a relatively low computational cost. Recall that we implemented the control points strategy with $N = 4$ points, where giving-up is myopically considered, located at equally spaced time intervals (corresponding to the horizontal yellow lines in Fig. 7). In comparison, the myopic threshold has $N \rightarrow \infty$ points, where giving up is considered and the play-to-win can be implemented with $N = 1$ point (at the very beginning). By channeling dropouts to a few control points, there is a substantial reduction in the probability of early dropouts. As seen in Fig. 7, the control points strategy not only approximates the optimal strategy in terms of rewards, but also in terms of the distribution of dropouts. While dropouts following the control points strategy only occur at distinct points, the relative mass resembles that of the optimal strategy much more than the other two boundedly rational strategies (e.g., the distributions for $\lambda_m = 5$ and $\sigma_m = 5, 10$). In high uncertainty environments where the decision about whether to begin the task is marginal, the control points strategy leads to

substantial gains over the myopic and play-to-win strategies (see $\lambda_m = 4$ and $\sigma_m = 10$ bottom left of Fig. 7), suggesting that it may be a good heuristic in high uncertainty environments. This implementation is not always beneficial, for instance, in environments with lower levels of uncertainty an additional control point at $t = 0$ would protect decision-makers from starting the make-or-break task when it is too challenging (upper left panel $\lambda_m = 4$ and $\sigma_m = 5$).

3.2.4. Discussion

Our results suggest that even though there is a well-defined optimal allocation strategy, the computational costs may be prohibitively expensive for both humans and artificial decision makers. However, we show that simpler boundedly rational alternatives can achieve high expected rewards at low computational costs by disregarding information and exerting less control. The myopic strategy produces a distinct pattern of risk-averse behavior (compared to the optimal strategy) at a lower computational cost. Using even simpler models, we show that the play-to-win and control points strategies largely outperform the myopic threshold and can sometimes approximate the performance of the optimal strategy. The competitive performance of these heuristic strategies runs contrary to findings from other complex decision-making settings, where it has been shown that higher uncertainty necessitates more monitoring and control (Osman, 2010). Our results suggest that the right level of complexity and control may depend on the structure of the task.

4. General discussion

How should people invest time in risky but highly rewarding, make-or-break tasks, forgoing other activities with safe rewards? How much effort should a scientist invest towards applying for a prestigious grant? When should an entrepreneur give up on a high-risk yet potentially high-reward startup that is currently performing worse than expected? Researchers across disciplines in the social and behavioral sciences have alluded to the crucial factors involved in making these decisions (e.g., Eccles & Wigfield, 2002; Radner & Rothschild, 1975; Vroom, 1964), but the problem has previously evaded rigorous analysis. In this article, we present a formal framework that addresses this conceptual gap and can be applied to allocation problems in psychology, economics, behavioral biology, and artificial intelligence.

4.1. On the large impact of small changes

We found that the expected returns of the make-or-break tasks are sigmoid in shape, which implies that similar real-world effort allocation problems may harbor discontinuities in the best possible time allocation policy. Small changes in the parameters of the problem can shift the prescribed allocation policy from one extreme (i.e., investing nothing) to the other (i.e., allocating all available resources; see Kukla, 1972; Vancouver,

More, & Yoder, 2008). For instance, consider two students with very similar skill levels studying in the same competitive program. The optimal strategy may prescribe one student to drop out entirely, while the other should double down and invest all available resources into their studies. Similarly, relaxing the deadline for a grant application or reducing the threshold required to achieve a bonus at work may mobilize people to dedicate themselves fully to a task, whereas they might have appeared indifferent before. Crucially, we showed that the optimal allocation policy is moderated by the degree of uncertainty in the relation between effort and performance. In environments with high uncertainty, people with a larger range of skill levels will be motivated to pursue challenging goals, either in hope of an outside chance of success or working hard to minimize the offhand chances of failure. These results are at odds with the prevailing theories of time allocation in economics, behavioral biology, and psychology, which typically assume diminishing returns or that different activities complement each other (e.g., Becker, 1965; Borjas & Van Ours, 2000; Charnov, 1976; Kurzban et al., 2013). In these theories, the underlying optimization problems are convex, with only a single optimum that shifts gradually as the parameters of the problem are varied. Likewise, the predicted policy switches cannot be accounted for by theories of motivation in cognitive and organizational psychology (Bandura & Locke, 2003; Heath, Larrick, & Wu, 1999; Locke & Latham, 2002), which suggest that people tend to respond to marginal changes in the difficulty of achieving a goal with continuous (either monotonic or non-monotonic) changes in the amount of effort exerted.

4.2. The computational challenges of dynamic decision-making

There are two modes of discovery in the study of human decision-making. Psychologists and neuroscientists tend to start their inquiry by conducting experiments reflecting realistic decision-making settings and assess the extent to which human behavior corresponds to plausible cognitive models (e.g., Brehmer, 1992; Busemeyer, 2002; Gonzalez, Lerch, & Lebiere, 2003). The descriptive, rather than the normative value of these models drives the inquiry. For example, mathematical psychologists and neuroscientists have extensively investigated dynamic decision-making in problems where people choose between two or more multi-attribute options, and where information about the value of the options is accessed and evaluated dynamically (e.g., by retrieving memories of similar items experienced in the past). Similarly to our setting, new information is assumed to unfold dynamically as a stochastic drift over time. Furthermore, these models assume that decision makers make a choice when the evidence accumulated in favor of an alternative has crossed a decision threshold (e.g., Busemeyer & Townsend, 1993; Khodadadi, Fakhari, & Busemeyer, 2017; Krajbich, Armel, & Rangel, 2010; Ratcliff et al., 2016; Usher & McClelland, 2001).

One of the main advantages of these dynamic models is that they generate rich predictions, not only about eventual choices, but also about the timing of events, namely, when and how often different decision thresholds will be crossed at different parameterizations of the problem. At the same time, it has been increasingly valuable to leverage dynamic

optimization techniques to obtain insights in the properties of optimal policies and to assess the conditions under which different action thresholds lead to good decisions (see Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Fudenberg, Strack, & Strzalecki, 2015; Malhotra et al., 2018; Tajima et al., 2016). These recent investigations provide the first steps towards a rational analysis (Anderson, 1991; Chater & Oaksford, 1999) or what David Marr (1982) describes as a computational level analysis of the tasks in question, developing a better understanding of the reward structures and the environment within which cognition adapts. Are the optimal policies for this task also cognitively plausible strategies? Our computational complexity analysis suggests it is unlikely. Nevertheless, knowing the optimal policies can guide the search for computationally efficient and psychologically plausible algorithms that can produce competitive performance.

In contrast, economists and scientists in neighboring disciplines begin their inquiry from normative models. Researchers in economics and finance have extensively studied optimal policies in dynamic decision-making settings where the value of an asset—commonly described by a Wiener process—fluctuates over time (e.g., a stock, or the value of a factory) and agents are required to make high-impact and irreversible decisions (e.g., selling a factory; see McDonald & Siegel, 1986; Pindyck, 1991). These problems have clear-cut analogs in common problems encountered by laypeople, such as whether and when to sell a house or quit a stressful job. Dixit and Pindyck (1994) were well aware of the relevance of their modeling framework for everyday decisions and briefly discuss marriage as an irreversible dynamic decision that most people make at least (or at most) once. To date, only a handful of experimental studies have investigated how people make decisions in such dynamic settings, where the tasks are framed as investment decisions (Oprea, Friedman, & Anderson, 2009; Strack & Viefers, 2013). How do people cope with the computational complexity of dynamic decision-making in these settings when the optimal strategy is computationally inaccessible and there are scant opportunities for learning? Defining and studying computationally simpler, boundedly rational strategies could lead to a rich set of predictions about behavior in dynamic decision environments (see, e.g., Oprea et al., 2009).

The dynamic task we present here has structural and technical similarities with both these strands of research. Thinking like economists or theoretical biologists, we were able to characterize the optimal giving-up strategy for a family of common problems.⁶ In the spirit of Herbert Simon (1978), we employed complexity theory—an essential methodological tool in computer science and a artificial intelligence—to gauge the computational requirements of the optimal strategy. We showed that it is computationally excessive for humans, and that it is demanding even for modern computers. Thinking like behavioral scientists, we then proposed several cognitively plausible alternatives, motivated by the principles of myopic and heuristic decision-making. These simpler strategies come with substantially cheaper computational costs. Deriving the optimal strategy enabled us to assess the prescriptive value of the boundedly rational strategies for different parameters of the decision-making problem. As in the dynamic decision-making problems investigated in mathematical psychology and neuroscience, our models generate a rich set of predictions about the timing of success and dropout at different parameterizations of the

problem. Clearly, we have not yet exhausted the space of psychologically plausible stopping strategies (see Busemeyer & Rapoport, 1988). For instance, an agent may decide to stop when performance stagnates for a long period of time, constructing a local (in time) stopping rule. Formally analyzing additional strategies is beyond the scope of this paper, yet our framework paves the way for future experiments that will make it possible to assess the descriptive value of the proposed strategies and to identify other plausible psychological alternatives.

4.3. Risky choices, sunk costs, and bounded rationality

We started from the assumption that decision makers maximize expected value. This was a convenience assumption that allowed us to convey our results with the greatest clarity. How do our results connect to the literature on risky choice? When the marginal utility of rewards diminishes (Bernoulli, 1954), people may discount the expected rewards from the two activities, especially from the large make-or-break bonus (see Supporting Information Section 6).⁷ Going a step further, people may subjectively reassess the value of succeeding or dropping out in the make-or-break task (Atkinson, 1957; Heath et al., 1999; Lewin et al., 1944) or distort the probabilities of success and failure at each level of invested effort (Kahneman & Tversky, 1979). How do these findings transfer to the complex and dynamic decision-making settings encountered in the real world? There is evidence that as the environment becomes more complex (e.g., the number of alternatives increases) people tend to rely more on computationally simple strategies (e.g., Venkatraman, Payne, & Huettel, 2014). Every variable transformation, such as those implied by expected and non-expected utility theories, requires computation. However, we have already shown that the optimal strategy is prohibitively expensive, regardless of the mode of computation. Thus, deriving an “optimal” risk-averse or time-discounted policy using backwards induction may have normative value (e.g., Smith & Ulu, 2017) but little descriptive or prescriptive value. Plausible theories of risky choice in dynamic contexts have to address the severe computational challenges of deriving optimal and time-consistent risk-averse strategies (also see Hammond, 1976). A way to tackle this issue is to assume that some people are described by a non-expected utility theory (e.g. prospect theory), while at the same time taking risks myopically, as if they disregarded future choices (Ebert & Strack, 2015; Thaler et al., 1997).

In the real world, people may make risky choices by responding to the reward structure of the environment learning, for instance, from experienced rewards (Denrell, 2007; Houston & McNamara, 1999; March, 1996; Stewart, Chater, & Brown, 2006). Importantly, computational limitations require that decision makers find simple and robust strategies that trade off effectively between complexity and expected returns. This argument has been advanced already for static risky choices (see Brandstätter, Gigerenzer, & Hertwig, 2006; J. W. Payne, Bettman, & Johnson, 1993), yet it is imperative in most dynamic decision-making problems, where the costs of computing “optimal strategies” are high even for artificial agents. The psychologically plausible strategies we have proposed can produce different risk-taking patterns, highlighting the boundedly rational

approach to the study of risky behavior in dynamic contexts. For instance, the myopic giving-up strategy produces risk-averse behavior when compared to the optimal strategy, and becomes increasingly risk-averse with higher levels of uncertainty (see Fig. 6). Risk-aversion could also be derived at much higher computational costs by passing rewards through a risk-averse utility function and then calculating the optimal strategy. However, here the pattern of risk-aversion produced by the myopic strategy is due to computational simplifications rather than some innate risk preference. The play-to-win strategy, by contrast, produces behavior that can be seen as a version of risk-seeking and is compatible with the sunk cost fallacy (Arkes & Blumer, 1985). Thus, instead of subjectively transforming rewards and probabilities, people might express their risk tendencies by committing to boundedly rational strategies (Friedman et al., 2014; Pedroni et al., 2017). Subjective transformations of value and distortions of perceived probabilities might be important components of dynamic choices, but they have to operate in tandem with other behavioral assumptions—such as myopia—to be behaviorally plausible (also see Thaler et al., 1997). We intend to investigate these issues in future experiments.

4.4. *Self-efficacy and learning*

When people embark on new, challenging projects (such as pursuing a PhD or founding a startup), they often only have a rough estimate of their own abilities. Bandura (1977) coined the term “self-efficacy” to describe the belief in one’s ability to succeed, complete tasks, and reach goals. A mismatch between believed and actual skill levels can alter how people allocate their time to different tasks. Under-confidence can have the most dramatic effect, leading to self-fulfilling prophecies, where people who underestimate their abilities may never pursue a highly rewarding make-or-break goal (see Hogarth & Karelaia, 2012). Without direct experience of successful achievement of a challenging goal, the decision maker may never rectify their initial beliefs. In the one-shot version of the task, overconfidence can lead decision makers to under-invest or over-invest depending on their skill level. In the dynamic version of the task, however, overconfidence may encourage myopic individuals to persevere for longer, thereby counteracting the conservative effect of myopic giving up (Nozick, 1994).

How do people form beliefs about their own skill level and the relationship between effort, luck, and achievement in a task? They may use social comparisons or individual experiences in similar tasks to inform their beliefs about their skill levels (Bandura, 1977; Bol et al., 2018; Frank, 1935). Repeated interactions with a task or continuous feedback about performance may help them to hone their intuitions about how effort and luck jointly define the chances of achieving a goal (Weiner, 1985; Weiner & Kukla, 1970). Even if people are certain about their skill level, they may leverage repeated interactions with the same task to refine their allocation strategies or choose among them (Erev & Barron, 2005; Rieskamp & Otto, 2006). Oprea et al. (2009), for instance, studied how people behave in optimal stopping problems, where they have to choose when to sell an asset. Although the optimal strategy is similar in terms of complexity to the optimal

threshold strategy described here, through repeat interactions with the same task, people were able to approximate optimality by leveraging a simple learning strategy.

4.5. The value and cognitive underpinnings of perseverance

A main insight from the optimal giving-up strategy we advanced is that optimally behaving agents should persevere before dropping out, and that the strategy should become more tolerant as the amount of uncertainty in the environment increases. This result echoes findings from the dynamic investing literature in economics and finance, where agents should wait longer before making irreversible decisions, thus harnessing the value of dynamically unfolding information (Dixit & Pindyck, 1994). Duckworth and collaborators have shown that the ability to persevere and maintain effort in the face of failures is at least as important as intelligence for succeeding in one's field (Duckworth, Peterson, Matthews, & Kelly, 2007). The unexpected outcomes of real-world problems can often lead to discouragement and make people doubt their chances of success. Perseverance when pursuing high-stakes, yet challenging goals might be even more valuable in real-world settings, precisely because of the additional uncertainties involved in estimating one's abilities. In such settings, simple strategies that persevere for longer than what seems reasonable under a myopic set of beliefs (e.g., the control points strategy that considers giving up at only a few discrete time points, or play-to-win as an extreme case of ignoring feedback) could in fact lead to better outcomes. In line with work by McGuire and Kable (2013), our approach shows that people's ability to persevere is not merely a character trait but has important cognitive underpinnings.

5. Conclusion

A common movie plot involves a protagonist pursuing a challenging but rewarding make-or-break goal. After multiple frustrations and drawbacks, the protagonist manages to come back on top before time runs out and marginally succeeds in meeting their goal. In real life, people may rationally abstain from investing in such risky make-or-break goals or give up early after initial setbacks, in order to pursue other safer activities. Knowing when to pursue a challenging goal and when to give up are difficult decision-making problems, yet they are pervasive in everyday life. We have shown that simple strategies relying on myopic or heuristic principles perform competitively against the optimal strategy, yet sidestep much of the computational burden. What is more, these boundedly rational strategies can reproduce the distinctive behavioral patterns of risk-aversion, risk-seeking, and the sunk-cost fallacy, suggesting that these patterns arise not only from inherent preferences, but could also be the product of simplifying cognitive strategies. Thus, the persevering movie hero may appear unboundedly courageous, exactly because they are boundedly rational.

Acknowledgments

We thank Daniel Barkoczi, Jerome Busemeyer, Michalis Drouvelis, Daniel Friedman, Siragan Gailus, Lisheng He, Chien-Ju Ho, Thorsten Joachims, Konstantinos Katsikopoulos, Emmanuel Kemel, Bobby Kleinberg, Amit Kothiyal, Thorsten Pachur, Rajiv Sethi, Lisa Son, Leonidas Spiliopoulos, Philipp Strack, Ryutaro Uchiyama, the participants of the Concepts and Categories seminar at New York University, and the members of the Judgment and Decision-Making laboratory at Indiana University for their insightful comments. We are grateful to Susannah Goss for editing the manuscript. This research was supported in part through NSF Award IIS-1513692.

Code availability statement

The code for producing the results reported in the paper is available at <https://osf.io/p2vur/>.

Notes

1. Before the introduction of formal decision theory (von Neumann & Morgenstern, 2012) to psychology by Edwards (1953, 1954a), there were several attempts by psychologists to introduce a formal theory of valence (Meehl & MacCorquodale, 1992); many of these attempts were inspired by the study of decision problems characterized by success–failure reward structures. This body of research, commonly referred to as expectancy theory, turned out to be quasi-equivalent to subjective expected utility theory (Feather, 1959; Siegel, 2007), but lacked the theoretical coherence and elegance of the latter, and waned in influence over time.
2. Lewin et al. (1936) and Atkinson (1957) suggest that failure to reach a previously set aspiration level is accompanied by negative payoffs (e.g., feeling embarrassment or disappointment). Similarly, expected utility theory suggests that people may discount rewards B . Even if we allow for subjective transformations of the experienced outcomes, the main insights of our framework hold.
3. There are various types of optimal stopping problems (for examples, see DeGroot, 2005). The dynamic decision-making problem we study shares concepts and methods with three distinct lineages of such problems. The first can be traced to Wald's sequential analysis (1964), where an agent has to decide when to stop evaluating alternative hypotheses. This line of thought has been pursued further in psychology and neuroscience, leading to choice models where evidence in favor of different alternatives unfolds dynamically in time as a stochastic process (Ratcliff, Smith, Brown, & McKoon, 1975; Tajima, Drugowitsch, & Pouget, 2013). A second lineage studied in economics and finance builds on stochastic processes to investigate

when to make high-stakes decisions when the value of assets fluctuate (Dixit & Pindyck, 1994; Jacka, 2015). The main theoretical result of this line of research is that agents should wait longer before making consequential decisions. A third strand of research in management science and operations research relies on dynamic optimization techniques to explore when to adopt a new technology or give up investing in it when new information about its potential unfolds dynamically (McCardle, 1982; Ulu & Smith, 1974).

4. The Fast-Fourier-Transform algorithm is considered one of the most important algorithms of the 20th century (Cipra, 2000) because it reduced the computational cost of the discrete Fourier transform from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$.
5. In the absence of uncertainty ($\sigma = 0$), the problem degenerates and both strategies follow the same solution as the one-shot problem, where the optimal allocation can be derived from the outset and the notion of a giving-up threshold becomes meaningless. That's exactly the version of the problem studied by Kukla (2010).
6. Note that it is not always feasible to derive the optimal strategy for dynamic decision-making problems. In multi-armed bandit problems, for example, the optimal policy (i.e., Gittins index) can only be derived under restrictive assumptions (Whittle, 1970). Similarly, psychologists have long investigated human decision-making in challenging dynamic tasks where it is impossible to even derive the optimal policy (Brehmer, 1992; Gonzalez et al., 2017).
7. Note that some form of intertemporal discounting (Berns, Laibson, & Loewenstein, 2007) is often built into dynamic decision-making models. We avoided adding a discount factor in the main text to prevent the mathematics from appearing more daunting without adding much substance. We explore that version of the problem in Supporting Information Section 5.

References

- Ainslie, G. (1992). *Picoeconomics: The strategic interaction of successive motivational states within the person*. New York: Cambridge University Press.
- Analytis, P. P., Kothiyal, A., & Katsikopoulos, K. V. (2014). Multi-attribute utility models as cognitive search engines. *Judgment and Decision Making*, 9(5), 403–419.
- Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences*, 14(3), 471–485.
- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1), 124–140.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64(6), 359–372.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215.
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87–99.
- Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, 75(299), 493–517.
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1), 23–36.
- Berns, G. S., Laibson, D., & Loewenstein, G. (2007). Intertemporal choice: Toward an integrative framework. *Trends in Cognitive Sciences*, 11(11), 482–488.

- 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), 700–765.
- Bol, T., de Vaan, M., & van de Rijt, A. (2018). The matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19), 4887–4890.
- Borjas, G. J., & Van Ours, J. C. (2000). *Labor economics*. Vol. 2. Boston, MA: McGraw-Hill.
- Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge, UK: Cambridge University Press.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113(2), 409–432.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81(3), 211–241.
- Brent, R. P. (1976). The complexity of multiple-precision arithmetic. In R. S. Anderssen & R. P. Brent (Eds.), *The complexity of computational problem solving* (pp. 126–165). Brisbane, Australia: University of Queensland Press.
- Busemeyer, J. (2002). Dynamic decision making. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 3903–3908). Oxford: Elsevier Press.
- Busemeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology*, 32(2), 91–134.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100(3), 432–459.
- Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9(2), 129–136.
- Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, 3(2), 57–65.
- Cipra, B. A. (2000). The best of the 20th century: Editors name top 10 algorithms. *SIAM News*, 33(4), 1–2.
- DeGroot, M. H. (2005). *Optimal statistical decisions*. (p. 82). Hoboken: John Wiley & Sons.
- Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review*, 114(1), 177.
- Diecidue, E., & Van De Ven, J. (2008). Aspiration level, probability of success and failure, and expected utility. *International Economic Review*, 49(2), 683–700.
- Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton, NJ: Princeton University Press.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101.
- Ebert, S., & Strack, P. (2015). Until the bitter end: On prospect theory in a dynamic context. *The American Economic Review*, 105(4), 1618–1633.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132.
- Edwards, W. (1953). Probability-preferences in gambling. *The American Journal of Psychology*, 66(3), 349–364.
- Edwards, W. (1954a). Probability-preferences among bets with differing expected values. *The American Journal of Psychology*, 67(1), 56–67.
- Edwards, W. (1954b). The theory of decision making. *Psychological Bulletin*, 51(4), 380–417.
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, 112(4), 912–931.
- Feather, N. T. (1959). Subjective probability and decision under uncertainty. *Psychological Review*, 66(3), 150–164.
- Fechner, H. B., Schooler, L. J., & Pachur, T. (2018). Cognitive costs of decision-making strategies: A resource demand decomposition analysis with a cognitive architecture. *Cognition*, 170, 102–122.
- Feldman, J. A., & Sproull, R. F. (1977). Decision theory and artificial intelligence 2: The hungry monkey. *Cognitive Science*, 1(2), 158–192.

- Festinger, L. (1942). A theoretical interpretation of shifts in level of aspiration. *Psychological Review*, 49(3), 235–250.
- Frank, J. D. (1935). The influence of the level of performance in one task on the level of aspiration in another. *Journal of Experimental Psychology*, 18(2), 159–171.
- Friedman, D., Isaac, R. M., James, D., & Sunder, S. (2014). *Risky curves: On the empirical failure of expected utility*. London: Routledge.
- Fudenberg, D., Strack, P., & Strzalecki, T. (2015). Stochastic choice and optimal sequential sampling. Available at arXiv:1505.03342v1.
- Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *The American Economic Review*, 96(4), 1043–1068.
- Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, 109(3), 416–422.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gonzalez, C., Fakhari, P., & Busemeyer, J. (2017). Dynamic decision making: Learning processes and new research directions. *Human Factors*, 59(5), 713–721.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4), 591–635.
- Hammond, P. J. (1976). Changing tastes and coherent dynamic choice. *The Review of Economic Studies*, 43(1), 159–173.
- Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, 109(1), 79–109.
- Heider, F. (1958). *The psychology of interpersonal relations*. Hillsdale: Lawrence Erlbaum Associates.
- Hogarth, R. M., & Karelaia, N. (2012). Entrepreneurial success and failure: Confidence and fallible judgment. *Organization Science*, 23(6), 1733–1747.
- Hoppe, F. (1931). Untersuchungen zur handlungs- und affektpsychologie. *Psychologische Forschung*, 14(1), 1–62.
- Hotaling, J. M., & Busemeyer, J. R. (2012). Dft-d: A cognitive-dynamical model of dynamic decision making. *Synthese*, 189(1), 67–80.
- Houston, A. I., & McNamara, J. M. (1999). *Models of adaptive behaviour: An approach based on state*. Cambridge: Cambridge University Press.
- Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J. ... Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 112(10), 3098–3103.
- Jacka, S. (1991). Optimal stopping and the American put. *Mathematical Finance*, 1(2), 1–14.
- Johnson, E., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory*, 104(1), 16–47.
- Johnson, E., & Payne, J. W. (1985). Effort and accuracy in choice. *Management Science*, 31(4), 395–414.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Khodadadi, A., Fakhari, P., & Busemeyer, J. R. (2017). Learning to allocate limited time to decisions with different expected outcomes. *Cognitive Psychology*, 95, 17–49.
- Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, 143(1), 131–141.
- Krajibich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Kukla, A. (1972). Foundations of an attributional theory of performance. *Psychological Review*, 79(6), 454–470.
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36(06), 661–679.

- Kushner, H., & Dupuis, P. G. (2013). *Numerical methods for stochastic control problems in continuous time*. New York: Springer Science & Business Media.
- Lewin, K. (1936). Psychology of success and failure. *Occupations: The Vocational Guidance Journal*, 14(9), 926–930.
- Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. M. Hunt (Ed.), *Personality and behavior disorders*. New York: Ronald Press.
- Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, 124(6), 762–794.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *The American Psychologist*, 57(9), 705–717.
- Malhotra, G., Leslie, D. S., Ludwig, C. J., & Bogacz, R. (2018). Time-varying decision boundaries: Insights from optimality analysis. *Psychonomic Bulletin & Review*, 25, 971–996.
- March, J. G. (1996). Learning to be risk averse. *Psychological Review*, 103(2), 309.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Henry Holt and Co.
- McCardle, K. (1985). Information acquisition and the adoption of new technology. *Management Science*, 31(11), 1372–1389.
- McCardle, K., Tsetlin, I., & Winkler, R. (2016). When to abandon a research project and search for a new one. INSEAD Working Paper No. 2016/10/DSC. Available at SSRN: <https://ssrn.com/abstract=2739699>.
- McDonald, R., & Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics*, 101(4), 707–727.
- McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review*, 120(2), 395.
- McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica*, 60(4), 803–836.
- Meehl, P. E., & MacCorquodale, K. (1951). Some methodological comments concerning expectancy theory. *Psychological Review*, 58(3), 230–233.
- Metcalfe, J., & Mischel, W. (1999). A hot/cold-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, 106(1), 3.
- Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Computational Biology*, 8(2), e1002342.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Nozick, R. (1994). *The nature of rationality*. Princeton, NJ: Princeton University Press.
- O'Donoghue, T., & Somerville, J. (2018). Modeling risk aversion in economics. *Journal of Economic Perspectives*, 32(2), 91–114.
- Oprea, R., Friedman, D., & Anderson, S. T. (2009). Learning to wait: A laboratory investigation. *The Review of Economic Studies*, 76(3), 1103–1124.
- Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological bulletin*, 136(1), 65.
- Papadimitriou, C. H. (2003). *Computational complexity*. Chichester, UK: John Wiley & Sons.
- Payne, J. W., Bettman, J., & Johnson, E. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.
- Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology: General*, 136(3), 370–388.
- Payne, J. W., Laughunn, D. J., & Crum, R. (1980). Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, 26(10), 1039–1060.
- Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, 1(11), 803.

- Peskir, G., & Shiryaev, A. (2006). *Optimal stopping and free-boundary problems*. Basel, Switzerland: Birkhäuser.
- Pindyck, R. S. (1991). Irreversibility, uncertainty, and investment. *Journal of Economic Literature*, 29(3), 1110–1148.
- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106(4), 643–675.
- Radner, R., & Rothschild, M. (1975). On the allocation of effort. *Journal of Economic Theory*, 10(3), 358–376.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281.
- Redner, S. (2001). *A guide to first-passage processes*. Cambridge University Press.
- Rieskamp, J., & Otto, P. E. (2006). Ssl: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135(2), 207–236.
- Rotter, J. B. (1942). Level of aspiration as a method of studying personality. A critical review of methodology. *Psychological Review*, 49(5), 463–474.
- Shiryaev, A. N. (2007). *Optimal stopping rules*. New York, NY: Springer Science & Business Media.
- Siegel, S. (1957). Level of aspiration and decision making. *Psychological Review*, 64(4), 253–262.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
- Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review*, 68(2), 1–16.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41(1), 1–20.
- Smith, J. E., & Ulu, C. (2017). Risk aversion, information acquisition, and technology adoption. *Operations Research*, 65(4), 1011–1028.
- Son, L. K., & Sethi, R. (2006). Metacognitive control and optimal learning. *Cognitive Science*, 30(4), 759–774.
- Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, 53(1), 1–26.
- Stojic, H., Analytis, P. P., & Speekenbrink, M. (2015). In D. C. Noelle, et al. (Eds.), *Human behavior in contextual multi-armed bandit problems* (pp. 2290–2295). Austin, TX: Cognitive Science Society.
- Strack, P., & Viefers, P. (2013). Too proud to stop: Regret in dynamic decisions. Available at SSRN: <https://ssrn.com/abstract=2465840>.
- Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature Communications*, 7, Article number: 12400.
- Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics*, 112(2), 647–661.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Ulu, C., & Smith, J. E. (2009). Uncertainty, information acquisition, and technology adoption. *Operations Research*, 57(3), 740–752.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592.
- Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, 32(6), 939–984.
- Vancouver, J. B., More, K. M., & Yoder, R. J. (2008). Self-efficacy and resource allocation: Support for a nonmonotonic, discontinuous model. *Journal of Applied Psychology*, 93(1), 35–47.
- Varian, H. R. (1978). *Microeconomic analysis*. New York: WW Norton.
- Venkatraman, V., Payne, J. W., & Huettel, S. A. (2014). An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence. *Organizational Behavior and Human Decision Processes*, 125(2), 73–87.
- Vroom, V. H. (1964). *Work and motivation*. New York: John Wiley & Sons.
- Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2), 117–186.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548.

- Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 15(1), 1.
- Whittle, P. (1980). Multi-armed bandits and the gittins index. *Journal of the Royal Statistical Society. Series B, Methodological*, 143–149.
- 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, 915–924.
- Zhang, S., & Yu, J. A. (2013). Forgetful Bayes and myopic planning: Human learning and decision-making in a bandit setting. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 26 (NIPS 2013)* (pp. 2607–2615). Red Hook, NY: Curran Associates.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. The optimal amount of time allocated to the make-or-break task as a function of the total available time $T \in [0, 10]$.

Figure S2. The process of backward induction that an agent has to follow to numerically calculate the returns from the optimal policy.