When to Keep Trying and When to Let Go: Benchmarking Optimal Quitting

Nikolay Sukhov*

Department of Physics, Princeton University

Rachit Dubey*

Department of Computer Science, Princeton University

Annie Duke

Department of Psychology, University of Pennsylvania

Thomas L. Griffiths

Department of Psychology and Computer Science, Princeton University

*Denotes equal contribution

Abstract

Persistence and perseverance, even in the face of great adversity, are admirable qualities. However, knowing *when to stop* pursuing something is as important as exerting effort toward attaining a goal. How do people decide when to persist and when to quit? Here, we design a novel task to study this question, in which people were given a finite number of opportunities to pursue stochastic rewards by selecting among a set of options that provide a reward each trial. At any time, if people were not satisfied with the option they had selected they could choose to abandon it and instead try a new option. However, if they did so they could never return to the previous option. Mathematical analysis of this task shows that the optimal strategy explores a relatively small number of options before settling on a sufficiently good option. Further, we find that the optimal strategy is to abandon an option if the total number of remaining trials exceeds a threshold specified by the observed option's performance. A large-scale, pre-registered experiment ($N = 3,632$) reveals that people largely behave in accordance with the optimal strategy. People also make decisions to persist with an option based on its performance and they typically explore relatively few options before settling on a sufficiently good one. However, compared to the optimal strategy, people are less sensitive to the number of remaining trials and are more likely to persist with sub-optimal options. Together, this work provides a new approach to studying how we decide when to quit and deepens our understanding of human persistence.

Keywords: perseverance, over-persistence, goal abandonment, quitting, rational model

When to Keep Trying and When to Let Go: Benchmarking Optimal Quitting

"Our greatest weakness lies in giving up. The most certain way to succeed is always to try just one more time."

– Thomas Edison

"If at first, you don't succeed, try, try again. Then quit. There's no use being a damn fool about it."

–W. C. Fields

GRIT and perseverance, even in the face of great adversity, are considered to be virtuous and commendable characteristics (Duckworth, Peterson, Matthews, & Kelly, 2007). Yet, no one goes through life without facing an insoluble problem. In such situations, over-persistence can waste time and resources, negatively affect mental health, and can be even self-defeating (Aspinwall & Richter, 1999; Åstebro, Jeffrey, & Adomdza, 2007; Baumeister & Scher, 1988; Halkjelsvik & Rise, 2015; Ordóñez, Schweitzer, Galinsky, & Bazerman, 2009; Wrosch, Scheier, & Miller, 2013). Indeed, knowing when to quit and stop persisting is as important as exerting effort towards attaining one's goal (Alaoui & Fons-Rosen, 2021; Beshears & Milkman, 2011; Wrosch, Scheier, Carver, & Schulz, 2003; Wrosch, Scheier, Miller, Schulz, & Carver, 2003).

However, deciding whether to quit or to keep trying is by no means an easy task. By continuing to pursue the current task, one could be missing out on other opportunities that might be even better. But, at the same time, success could be just around the corner, so perhaps it could be worth trying one more time. How much persistence is too much? How do people resolve this "pursue vs. quit" trade-off? While the field has made considerable strides towards understanding goal disengagement and over-persistence (McGuire & Kable, 2013; Milyavskaya & Werner, 2018; Wrosch, Scheier, Miller, et al., 2003), a systematic investigation of this trade-off can further improve our understanding of how people decide when to quit.

In this article, we present a formal analysis that benchmarks optimal quitting when pursuing options under uncertainty. We introduce a novel optimization task that captures the trade-off between pursuing an option vs. permanently abandoning it in the hoping of finding a better option in the future. Using this task, we specify how a rational agent should resolve the pursue vs. quit trade-off and investigate how people conform and deviate from optimal behavior. A key insight of our work is that the decision of whether to quit or persist with an option can be characterized using a mathematical optimization problem that combines two well-studied problems – the multi-armed bandit problem (Cohen, McClure, & Yu, 2007; Vermorel & Mohri, 2005) and the secretary problem (Ferguson, 1989; Seale & Rapoport, 1997).

In a standard multi-armed bandit problem, people have to choose between a set of options ("arms"), each with different unknown reward rates, to maximize the total reward they receive over a fixed number of trials (Gittins, 1979; Lai, Robbins, et al., 1985; Robbins, 1952). Multi-armed bandits are widely used within psychology and neuroscience for studying human decision-making under uncertainty, primarily because they require striking a balance between exploiting a reasonably well-known arm and exploring less-known arms for an even better alternative (Cohen et al., 2007; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Steyvers, Lee, & Wagenmakers, 2009). However, in the standard version of this task, the arms are *always available* to a decision-maker: even if an arm is abandoned, it is always possible to choose to return to that arm at a later point in time. Multi-armed bandits thus do not offer a direct way to study persistence and quitting.

In the secretary problem, people are presented with options sequentially and have to select the best option, with the caveat that they cannot return to an option once they have rejected it (Baumann, Schlegelmilch, & von Helversen, 2022; Baumann, Singmann, Gershman, & von Helversen, 2020; Freeman, 1983; Goldstein, McAfee, Suri, & Wright, 2020; Lee, 2006; Seale & Rapoport, 1997). These problems require resolving the trade-off between accepting a possibly sub-optimal option prematurely or rejecting it and hoping for a better option in the future. In standard versions of the secretary problem, the rewards provided by each option are fixed (and not stochastic), and so, as soon as a decision-maker is presented with an option, they can make the choice to either accept the option or abandon it. However, real decisions to continue a job or relationship have to take into account the inherent stochasticity in the rewards they provide (e.g., a job might be highly rewarding one day but then might not be so enjoyable on another day and so one needs to consider these variations when deciding whether to abandon the current job or look for another job). Therefore, while the secretary problem is ideal for studying *searching* under uncertainty, it does not fully capture the stochasticity that is a key part of decisions to persist or quit.

Here, we introduce the “pursue vs. quit” task, a novel optimization problem that combines the stochasticity of multi-armed bandit problems with the “no return” structure of the secretary problem. In our task, a decision-maker is provided with a series of options sequentially and has the goal of maximizing total rewards within their lifetime. Each option provides rewards with an initially unknown probability and the only way to learn the reward probability of an option is by exploring it. Unlike the standard multi-armed bandit task, once the decision-maker abandons an option, they can never go back to it. Further, the decision-maker is given only a finite number of trials overall to explore all the options. This puts a natural limit on how much one should explore any single option and so, at each time point, the decision-maker faces the difficult choice of either persisting with the option they have or permanently abandoning it in the hope of a better option in the future. This permanent abandonment aspect is similar to the secretary problem, but here the decision-maker is placed in an uncertain environment where the quality of the option has to be inferred through exploration. Thus, our task allows us to systematically examine the trade-offs between persistence and quitting.

The pursue vs. quit task has a precise solution that maximizes the expected return and that can be found analytically. The optimal strategy gives deeper insight into the nature of the problem and serves as a benchmark for studying human decision-making. We derive this optimal strategy and find that it explores a relatively small number of options before settling on a sufficiently good option and exploiting it for the remaining trials. Further, the optimal strategy prescribes abandoning an option only when the number of remaining trials exceeds a threshold that is specified by the observed option’s performance. To explore how people’s behavior compares to the optimal strategy, we conduct a large-scale, pre-registered experiment ($N = 3632$) and find that people largely follow the optimal strategy in transitioning from exploration to exploitation. However, people are significantly more risk-averse and tend to keep pursuing poor options that the optimal strategy indicates to abandon. Further, the dependence on the number of remaining trials, although statistically significant, is much smaller in magnitude, and reverses in direction i.e., people are more likely to quit a poor option toward the end of their run.

Open Practices Statement

In the following sections, we present the results of a computational study followed by the results of a large-scale empirical study. The code and data for the model simulations can be accessed here: https://github.com/sukhovn/OptimalQuittingExperiment/. Prior to collecting the behavioral data, we pre-registered the study on OSF. The pre-registration included the data collection protocol and the data analysis plan (`https://osf.io/3w7bx`). The behavioral data and analysis code of this data is publicly available at `https://github.com/sukhovn/OptimalQuittingExperiment/blob/master/ExperimentProcessing.ipynb`.

The pursue vs. quit task

To study the pursuit and abandonment of options under uncertainty, we design a novel sequential decision-making task where options provide rewards stochastically and are no longer available if one chooses to forego them (Fig 1). In what follows, we introduce the task using the language of the multi-armed bandit task, in which options are referred to as "arms" that the participant "pulls". Formally, the decision-maker is provided with a set of N Bernoulli arms and T opportunities to pull them ($N \gg T$). Upon pulling an arm, the arm produces with probability θ, $0 \leq \theta \leq 1$, reward $r = 1$ (which we term a "success"), and with probability $1-\theta$, reward $r = 0$ (which we term a "failure"). The probability θ is *a priori* unknown and different for each arm. The goal of the decision-maker is to maximize cumulative reward, $\sum_{t=1}^{T} r_t$.

In the beginning, the decision-maker is provided with an arm sampled randomly from the series of arms. At each time point, the decision-maker can either pull the current arm and observe the reward outcome (which decreases their remaining pulls by 1) or they can choose to quit the current arm (which doesn't affect their remaining pulls). If the decision-maker decides to quit the current arm, they are provided with a new arm (sampled randomly without replacement) and the process repeats until the decision-maker exhausts their finite number of pulls. This task, which we name the "pursue vs. quit" task, is thus a hybrid of the multi-armed bandit task (where options provide rewards stochastically) and the secretary problem (where options are no longer available if one chooses to forego them) and allows us to systematically study the pursuit and abandonment of options.

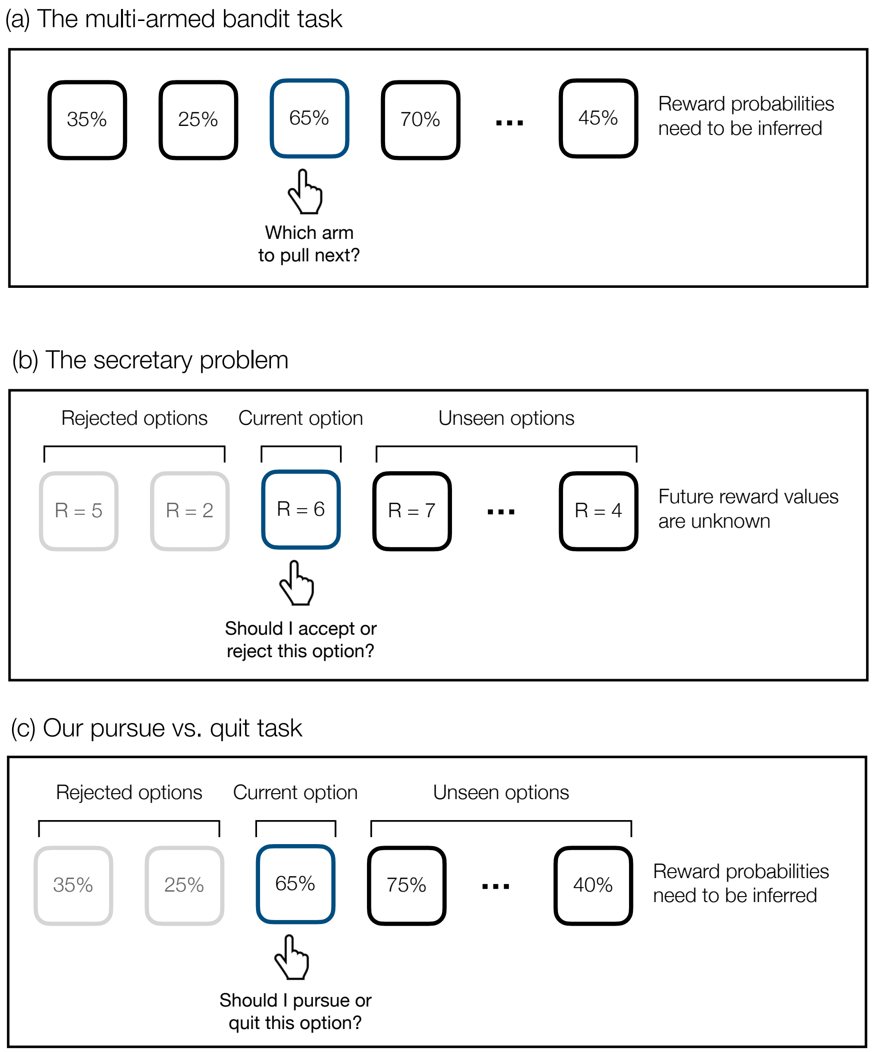

Figure 1. **Illustration of our task** (a) In a multi-armed bandit task, one has to choose between a set of options, each with different unknown reward rates. Here, the arms are always available to the decision-maker. (b) In the secretary problem, one has to select the best option out of a series of options and cannot return to an option one has rejected. Here, the rewards provided by each option are fixed (and not stochastic). (c) In our pursue vs. quit task, each option has a different unknown reward rate and once the decision-maker rejects an option, they cannot return back to it.

How an optimal agent decides to pursue or quit options

The pursue vs. quit task has an optimal strategy that maximizes the expected return and that can be determined analytically. In this section, we outline the key features of this optimal strategy to provide insights into how a rational agent resolves the pursue vs. quit trade-off. Interested readers are referred to

Table 1

The required threshold for the number of remaining pulls before the optimal strategy suggests quitting the current arm.

Quitting threshold for the number of remaining pulls									
# of Failures / # of Success	1	2	3	4	5	6	7	8	9
0	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
2	12	1	1	1	1	1	1	1	1
3	39	5	1	1	1	1	1	1	1
4	94	10	3	1	1	1	1	1	1
5	190	19	6	3	1	1	1	1	1
6	344	30	10	5	2	1	1	1	1
7	578	45	14	7	4	2	1	1	1
8	915	63	20	9	5	3	2	1	1
9	1385	86	27	13	7	5	3	2	1

Note: Various combinations of successes and failures are shown. For example, when the decision-maker has observed 9 successes and 1 failures, then the decision-maker should only quit when the number of remaining pulls is greater than or equal to 1385 (bottom row). However, having observed 9 successes and 8 failures, they should quit the arm if the number of remaining pulls is greater or equal to 2.

the Appendix for the mathematical details of the derivation of the optimal strategy. The code and data for the model simulations can be accessed here: https://github.com/sukhovn/OptimalQuittingExperiment/.

Feature 1: Always persist with a great option. The first feature of the optimal strategy concerns great options. If the current arm has only resulted in successes so far (and no failures), then the optimal strategy recommends persisting with that arm: one should never leave an option that has only given successes and no failures.

Feature 2: Always abandon a bad option immediately. The second feature concerns what an optimal agent should do with bad options. At any time, if the number of failures of the current arm is greater or equal to the number of successes, then the optimal strategy recommends immediately quitting that arm and moving to the next arm in order to maximize total rewards. Thus, the optimal agent never persists with an option where the successes don't exceed the failures.

Feature 3: Consider how much time is left when considering good options. What should a decision-maker do when an option is not outright great or bad? According to the optimal strategy, if an arm has resulted in some failures but the number of successes is greater than the number of failures, then one should take into account how much time one has left. For each such option, there exists a threshold τ such that if the number of *remaining* pulls is greater or equal to τ, then the optimal strategy recommends quitting. However, if the number of remaining pulls is less than τ, then it is optimal to persist with the current option in order to maximize total rewards. Thus, the decision to quit depends not only on how good the current option is but also on how many pulls the decision-maker has remaining.

As an illustration of this feature, Table 1 shows the required number of remaining pulls before the optimal strategy suggests quitting the current arm for various potential combinations of successes and failures. If the current arm is very good (e.g., when it has given 9 successes and only 1 failure), then the value of the threshold is very large (= 1385). However, if the current arm is not exceptional (e.g., it has given 9 successes and 8 failures), then this threshold is quite small (= 2). Thus, the optimal strategy suggests quitting a very good arm only if the number of remaining pulls is very large (so the chance of finding an even better arm in the future is still high) but if the arm is not exceptional, then the optimal policy shows less patience and quits sooner.

Behavior of the optimal strategy

To illustrate the behavior of the optimal strategy, we ran the optimal strategy 10000 times on a series of arms with probabilities drawn from a uniform distribution (in each run, the decision-maker was given a total of 100 pulls). Figure 2(a) shows the histogram for the arms pulled by the optimal strategy, which explores a very small number of arms on average (average arms pulled ≈ 9). Figure 2(b) plots the contribution of individual arms. In the vast majority of runs ($> 98\%$), the last arm provided the leading contribution. Furthermore, the average total contribution of the arm with the largest contribution was around 85%. Thus, the optimal strategy explores a relatively small number of arms until it stops at one, which is then exploited for the remaining pulls. Further, that last arm provides almost all the reward

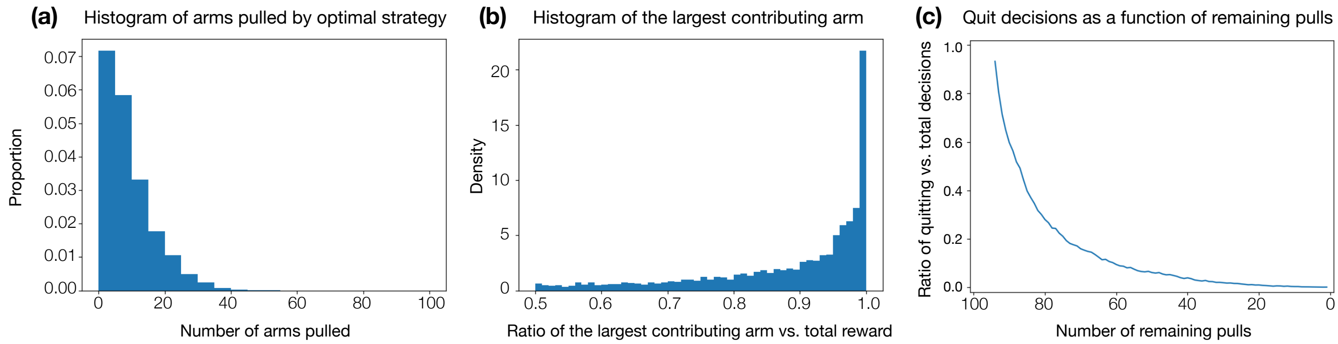

Figure 2. **Behavior of the optimal strategy.** (a) Distribution of the number of arms pulled by the optimal policy over 10000 runs (total pulls = 100 in each run). (b) Histogram of the contribution of the largest contributing arm. (c) The ratio of quit vs. total decisions as a function of remaining pulls.

accumulated in the optimal agent's lifetime.

Figure 2(c) shows, for all cases where the number of failures was at least 1, the ratio of quitting decisions against the total decisions as a function of the remaining number of pulls. Here, we see that the probability that the optimal strategy quits an arm is very high when the remaining pulls are high but then decreases as the number of remaining pulls becomes lower. This graph illustrates that the optimal strategy heavily considers the remaining number of pulls before quitting an option.

Behavioral Experiment: How do people decide whether to quit or pursue an option?

Having examined the behavior of the optimal strategy, we sought to study how people decide whether to quit or pursue options. To do so, we conducted a behavioral experiment in which participants took part in a sequential decision-making task that required them to decide whether to abandon an option in order to pursue a new one. Participants were provided with various options one by one in random order, where each option produced with some probability a reward, the probability of which was *a priori* unknown to the participant. At each time point, they could choose to either persist with their currently provided option (so they could keep getting the rewards associated with this option) or they could choose a different option instead (which provided a reward with a different unknown probability). Crucially, once they decided to forego their current option, that option became no longer available to them and they could then only interact with the other remaining options. Prior to collecting the data, we pre-registered the study on OSF. The pre-registration included the data collection protocol and the data analysis plan (`https://osf.io/3w7bx`). Data and analysis code is publicly available at `https://github.com/sukhovn/OptimalQuittingExperiment/blob/master/ExperimentProcessing.ipynb`.

Participants

We recruited 3632 participants from the online platform Prolific. They earned $1.8 for participation with the option to earn an additional bonus of up to $1. Unknown to the participants, and at the start of the experiment, participants were randomly assigned to three conditions: 0.9 ($N = 1214$), 0.8 ($N = 1203$), and 0.7 ($N = 1215$). In the 0.9 condition, the best button (out of all possible buttons that were available) provided a reward of 1 with a probability of 0.9. In the 0.8 condition, the best button provided a reward of 1 with a probability of 0.8. In the 0.7 condition, the best button provided a reward of 1 with a probability of 0.7.

Procedure

Participants took part in a task similar to the pursue vs. quit task described above (refer to Figure 3). Participants were provided a series of buttons to press and were given a finite number of button presses ($= 100$). Whenever they pressed a button, it provided with some (unknown) probability either 1 point or 0 points. Each button had a different chance of providing a point, e.g., a button could give a point 9 out of 10 times while another would only give a point 1 out of 10 times. Note that a particular button had the same chance of providing a point across presses. Participants were instructed that the probability with which each button would give them a point was randomly determined, and all probabilities were equally likely. That is, every time they selected a new button, they were equally likely to get a "good" or a "bad" button.

At the start of the experiment, participants were provided with a randomly selected button to press. At each time point, they could choose to either press their current button or they could request a new button. Requesting a button had no cost – they could request a new button as many times as they wanted. However, if they decided to move to the next button, they were not able to press any of their previous buttons, that is, they had to "abandon" the current button. Participants were paid a bonus of $0.01 per point and their goal was to get as many points as possible given the number of presses they had.

Participants were provided detailed instructions about the task, including the number of total button presses they had. After they were provided with the instructions (and before they began the task) they

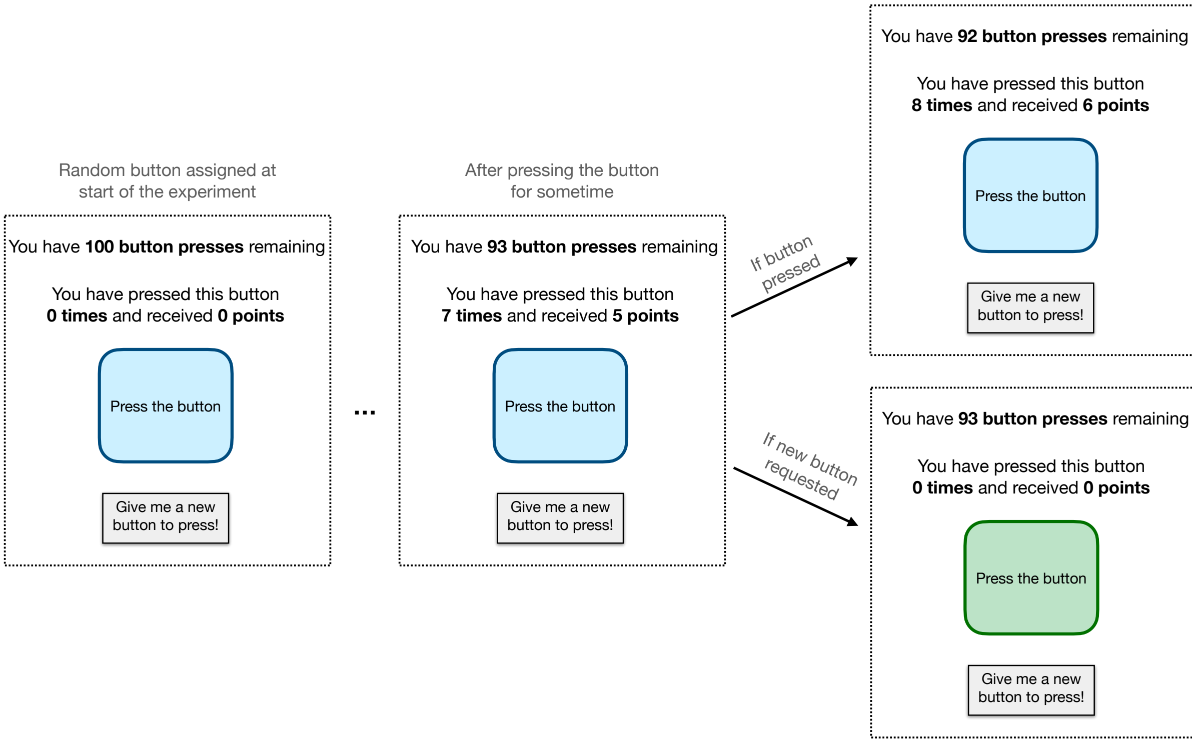

Figure 3. **Design of the experiment.** At the beginning of the experiment, participants were assigned a random button that gave them with some (unknown) probability either 1 point or 0 points. At each time point, they could decide to either press the existing button or they could request a new button (which would provide points with a different, unknown probability). Requesting a new button had no cost, however, it meant that they could no longer go back to the previous button.

were quizzed about the task and if they failed the quiz, they were asked to read the instructions again (this was repeated until they correctly answered all questions in the quiz). To familiarize participants with the task, before the main experiment participants were also given a practice run that had no restriction on the number of remaining presses and didn't count towards the monetary reward.

In order to avoid extreme scenarios where a button always resulted in success (and thus would never be abandoned), we capped the probability with which a button could provide points. To that end, participants were randomly assigned to three different conditions with button probability caps set to 0.7, 0.8, and 0.9. That is, in the 0.9 condition, the probability with each a button could provide rewards was drawn from a uniform distribution $U(0,0.9)$, and in the 0.8 condition, it was drawn from $U(0,0.8)$, and in the 0.7 condition, drawn from $U(0,0.7)$. This cap was unknown to the participants and it had no effect on their decision-making strategy (demonstrated in the next section).

Results

For all analyses that follow, we excluded participants who failed to answer the quiz correctly more than once as well as excluded the participants who kept pressing an obviously bad button for a very long period of time (we removed all participants who persisted with a button whose number of failures exceeded the number of successes by 20 or more). This led to the exclusion of 242 participants. The resulting numbers of participants in each condition were: $N = 1151$ (0.9 condition), $N = 1125$ (0.8 condition), and $N = 1114$ (0.7 condition).

We begin by looking at the general features of the decision-making strategies of participants. First, we find that participants overwhelmingly persisted with a button if it had given them no failures (probability of persisting > 0.97 when no failures have been observed in all three groups as seen in Table 2). This behavior is similar to Feature 1 of the optimal strategy which recommended always persisting with options that provided no failures ($z = 165.9$, $p < 0.001$ under a hypothesis that the decisions to persist were made at random).

However, compared to Feature 2 of the optimal strategy, participants were more likely to stay with a button even if it produced more failures than successes. Figure 4(a) shows participants' persistence for the buttons that resulted in more failures than successes (the optimal strategy would always abandon these buttons). When the difference between failures and successes was relatively small (e.g., 2), participants were significantly more likely to continue persisting with the button as compared to the optimal policy (e.g., close to 50% of buttons were persisted for all conditions when the difference was 2). If the number of failures exceeded the number of successes by a large margin, then participants behaved similarly to the optimal strategy with most of the participants quitting these bad buttons (e.g., 96.6% of buttons that had more failures than successes were abandoned before the number of failures exceeded successes by 5). Note that the plots were the same for the three cap groups, which is consistent with our assumption that participants use the same decision-making strategy irrespective of the probability cap. Together, this analysis suggests that compared to the optimal strategy, people are more reluctant to quit an option and their threshold to quit a bad option is much higher.

Next, we investigated the exploration patterns of participants. As seen in Figure 4(b), participants explored a relatively small number of buttons on average, which was similar to the behavior of the optimal

Table 2

Average characteristics of people's decision-making strategies in the three different probability cap conditions.

Probability truncation	0.9	0.8	0.7
Probability of persisting with a button when no failures have been observed	97.5%	97%	98%
Average number of explored buttons	7.7	8.9	10.7
Percentage of runs where last button's contribution is dominant	66%	57 %	51 %
The average total contribution of the arm with the largest contribution	69%	65%	58%

strategy (Figure 2(a)). Note that the average number of buttons explored increases as the probability threshold is lowered (also refer to Table 2). Next, for the majority of participants, the last button provided the leading contribution to the reward (refer to Table 2). This is also demonstrated in Figure 4(c) which shows that in the majority of cases, most points earned by the participants come from a single button. Thus, similar to the optimal strategy, people also explore a relatively small number of arms until they stop at one, which they usually pull until the end of the run and that last arm provides almost all the points accumulated during the run.

We next investigated how people's behavior compared to Feature 3 of the optimal policy. That is, we examined whether people's decision to quit also depended on how many presses they had remaining. Figure 4(d) considers all decisions made about buttons that had at least one failure and plots the ratio of quitting decisions against the total decisions as a function of the remaining number of presses. The probability that participants quit an arm was high when the remaining presses were high but decreased as the number of remaining presses became lower ($z_{0.9} = 9.9$, $p_{0.9} < 0.01$, $z_{0.8} = 9.1$, $p_{0.8} < 0.01$, $z_{0.7} = 9.2$, $p_{0.7} < 0.01$ that the decreasing trend is absent in Mann-Kendall test in each cap group respectively). Thus, consistent with the optimal strategy, people were less likely to quit as the number of remaining presses decreased.

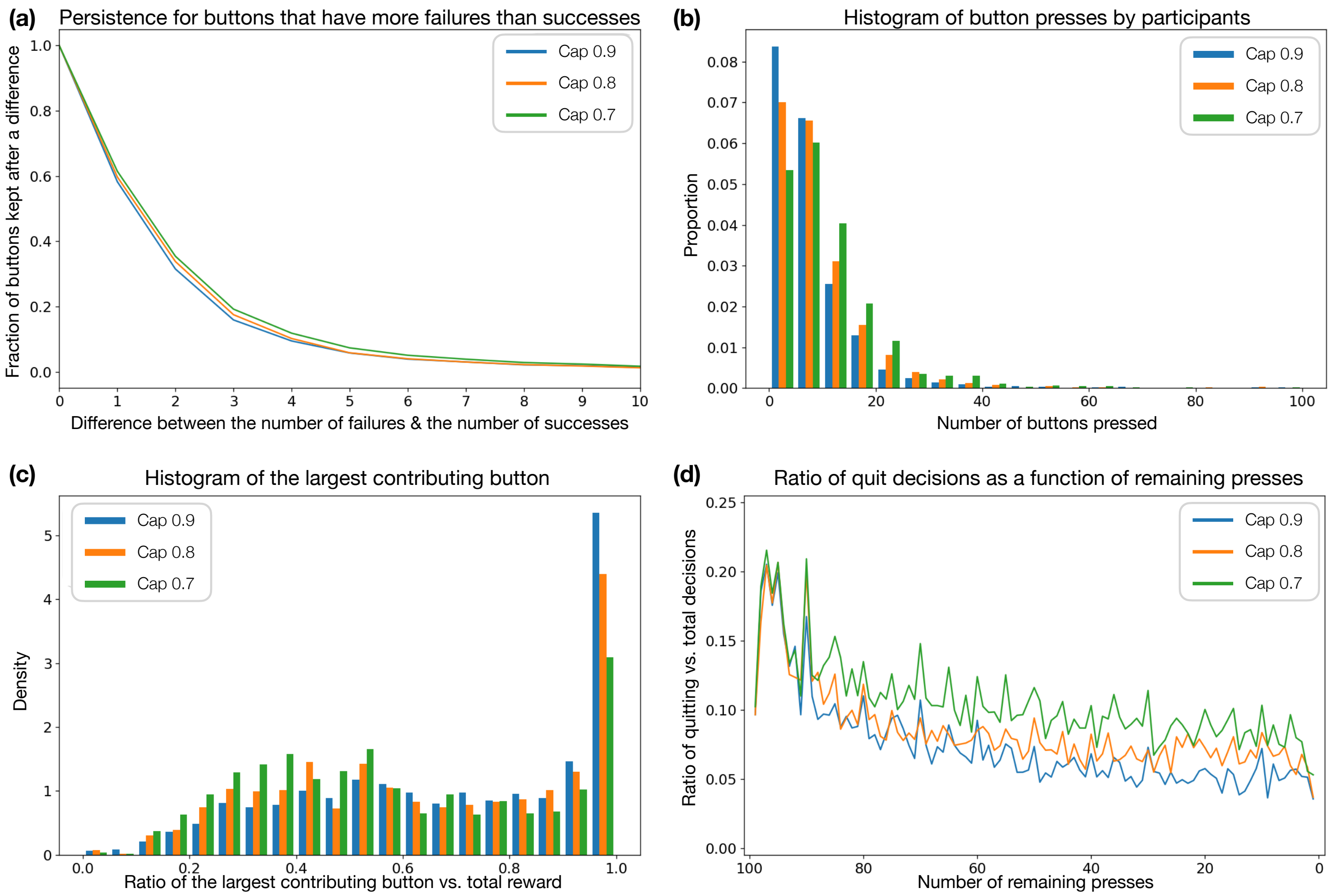

Figure 4. **Behavioral results.** (a) Fraction of buttons that have more failures than successes that were still kept after a certain difference between failures and successes. (b) Distribution of the number of buttons pressed by the participants. (c) Histogram of the contribution of the largest contributing button. (d) The ratio of quit vs. total decisions as a function of remaining presses.

We conclude our analysis by fitting the data (two classes corresponding to stay and quit decisions in 3D space spanned by the number of successes, failures, and remaining presses) with a single logistic regression model. First, we test the applicability of the logistic regression model by simulating 10000 runs performed by the optimal strategy. Figure 5(a) shows the receiver operating characteristic (ROC) curve for the logistic regression classifier trained on the simulated decisions made by the optimal strategy. The classifier has a very small error rate and thus logistic regression separates stay and quit decisions extremely well. Figure 5(b) provides more details of the regression analysis that was applied to the optimal strategy. The success rate parameter is -1.2, which indicates that the logistic regression model tends to stay more the more wins an arm has. The failure rate parameter is 2.9, which indicates that the

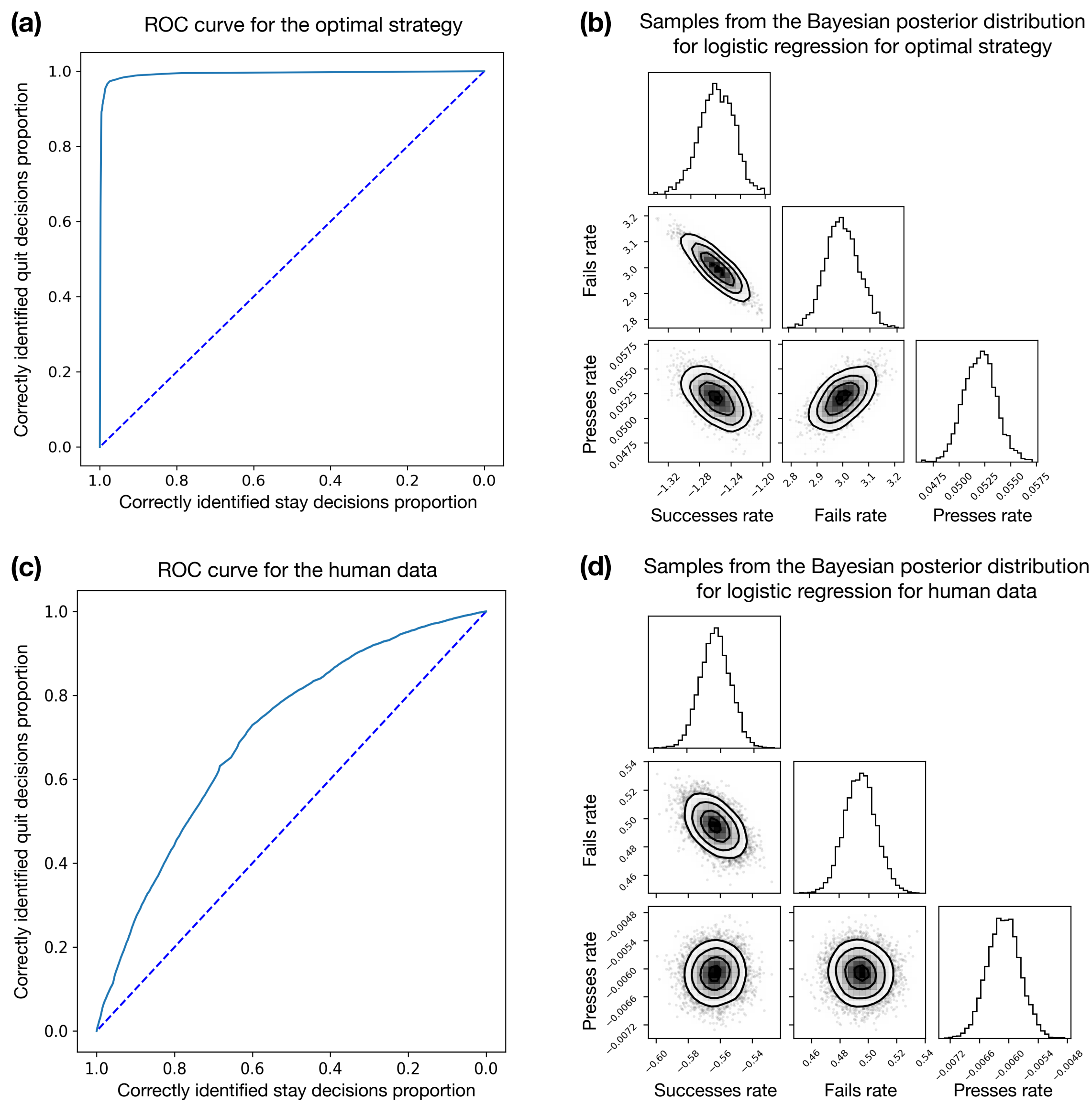

Figure 5. **Logistic regression fit to the optimal model and the behavioral data.** (a) The receiver operating characteristic (ROC) curve for the logistic regression classifier trained on the simulated decisions made by the optimal strategy. (b) Samples from the Bayesian posterior distribution for the logistic regression parameters for optimal strategy. (a) ROC curve for the classification of participants' responses. The curve is close to the random classifier line which indicates that the decision-making is highly probabilistic. (b) Samples from the Bayesian posterior distribution for the logistic regression parameters for human data.

more failures an arm has, the more the model tends to leave. The remaining press rate parameter is 0.05, which indicates that the model tends to stay more the fewer presses remain. All of these predictions are consistent with the full optimal strategy. Note that the logistic regression model intercept is -0.7.

Figure 5(c) plots the ROC curve for the logistic regression classifier trained on participants' responses. Note that we restricted ourselves to the group of decisions where the number of successes was between 1 and 5 and where the number of failures was between 1 and 4 (to examine how people made decisions about buttons that were neither outright good or bad). We found that the regression results were independent of the threshold probability (details of the logistic regression are provided in the Appendix). Further, the results also did not depend on the amount of time and the presses that participants spent in the practice round (we analyzed those participants that spent more than the average amount of time in the practice round and fitted their data with a separate logistic regression classifier and compared the regression parameters to the parameters of the logistic regression made for all participants and found no difference between the two groups; results of the analysis are provided in the Appendix).

Figure 5(d) provides more detail of the regression analysis. The success rate parameter is -0.56, 95% $CI[-0.58, -0.54]$, which indicates that participants tend to stay more the more successes they have, which is consistent with the optimal strategy ($\chi^2(1) = 5188$, $p < 0.01$ in favor of no wins dependence for the likelihood-ratio test). The failure rate parameter is 0.49, 95% $CI[-0.51, -0.47]$, which indicates that participants tend to leave more the more failures they have, which is also consistent with the optimal strategy ($\chi^2(1) = 2154$, $p < 0.01$ in favor of no fails dependence for the likelihood-ratio test). The unusual result comes with the number of remaining presses rate parameter which is -0.0061, 95% $CI[-0.0068, -0.0055]$, which indicates that participants tend to leave more the fewer remaining presses they have. The evidence for the effect is statistically significant ($\chi^2(1) = 339$, $p < 0.01$ in favor of no remaining presses dependence for the likelihood-ratio test), which contradicts the optimal strategy (in the model, we get a significant reverse dependence on the number of remaining presses). This indicates that when participants are considering buttons that are neither extremely good or bad, then they are more likely to leave them when they have fewer presses remaining (whereas the optimal strategy prescribes the exact opposite). Thus, compared to the optimal policy, participants are less likely to settle for moderate options toward the end of the run.

Discussion

Human grit can move mountains, but what should one do if the mountain refuses to move? How long should one keep trying before giving up? In this article, we introduced a novel task to study the difficult trade-off between pursuing an option vs. quitting. Using this task, we provided a mathematical analysis of how an optimal agent should resolve this dilemma as well as explored how people make decisions compared to this optimal strategy. Our mathematical analysis suggested that the optimal strategy is to always persist with an option that has resulted in no failures and to immediately abandon an option if the number of failures is greater or equal to the number of successes. Further, for options that are neither outright great nor bad, the decision-maker should always consider the remaining lifetime before making a decision. Finally, we found that the optimal strategy explores a few arms early in the lifetime before transitioning to the exploitation of a single option later in the lifetime i.e., optimal persistence requires abandoning some good options.

Similar to the optimal strategy, we found that people also tended to explore more during the early stages of the experiment before transitioning to exploitation in the later stages of the experiment. Yet, compared to the optimal strategy, people were more reluctant to quit the slightly poorer options. This provides some important insights into possible oversights people may make when they are deciding to pursue or quit options. People might forego seemingly “good” options early during their lives (e.g., a romantic partner or a job), but might be more likely to continue with those same options later during their lives. Yet, at the same time, people also have the tendency to stick with the not-so-good options longer than necessary, which can lead them to miss out on other better opportunities.

Even though our task had no explicit cost of switching, people persisted with sub-optimal options longer than prescribed, which opens up several questions for future research. For one, it would be important to examine the mechanisms that lead to over-persistence. People might be over-persisting because they might have an optimism bias (Koole & van’t Spijker, 2000; Tenney, Logg, & Moore, 2015; Zhang & Fishbach, 2010) or they might be pursuing those tasks because of sunk cost (Kelly, 2004; Sweis et al., 2018) or they might have simply been insensitive to the opportunity costs (Northcraft & Neale, 1986). Our work also connects to the literature on impulsivity and delay of gratification (McGuire & Kable, 2013; Mischel, Ayduk, & Mendoza-Denton, 2003; Schelling, 1984), and future work should investigate how self-control interacts with the decision-making in our task.

Another promising direction for future research would be to consider richer and more complex variations of our proposed task. For instance, it would be interesting to consider how a decision-maker should pursue and abandon options in non-stationary environments. How long should one stick with a seemingly poor option that has some (unknown) probability of providing higher rewards in the future? Relatedly, it would also be interesting to consider a variant of our task where different options are related to each other and so in this task, the decision-maker will have to learn from prior experience.

Everyday life presents us with many difficult situations where we are not sure whether to persist with something or abandon it (e.g., staying in a nice job vs. quitting it and hoping for an even better job in the future). By considering how an optimal agent should resolve this difficult pursue vs. quit trade-off and examining how people conform and deviate from the optimal strategy, our work provides a new set of tools for understanding the pursuit and abandonment of options. We hope that this work encourages further computational and empirical investigations of when people quit, including the many factors that influence these decisions.

Author Contributions

All authors developed the study concept. N. Sukhov wrote the software and performed the formal analysis with contributions from R. Dubey. T.L. Griffiths supervised the study design and model development. All authors discussed the results. N. Sukhov and R. Dubey drafted the manuscript, and A. Duke and T.L. Griffiths provided critical revisions. All authors approved the final version of the manuscript for submission.

References

Alaoui, L., & Fons-Rosen, C. (2021). Know when to fold'em: The flip side of grit. *European Economic Review*, *136*, 103736.

Aspinwall, L. G., & Richter, L. (1999). Optimism and self-mastery predict more rapid disengagement from unsolvable tasks in the presence of alternatives. *Motivation and Emotion*, *23*(3), 221–245.

Åstebro, T., Jeffrey, S. A., & Adomdza, G. K. (2007). Inventor perseverance after being told to quit: The role of cognitive biases. *Journal of behavioral decision making*, *20*(3), 253–272.

Baumann, C., Schlegelmilch, R., & von Helversen, B. (2022). Adaptive behavior in optimal sequential search. *Journal of Experimental Psychology: General*.

Baumann, C., Singmann, H., Gershman, S. J., & von Helversen, B. (2020). A linear threshold model for optimal stopping behavior. *Proceedings of the National Academy of Sciences*, *117*(23), 12750–12755.

Baumeister, R. F., & Scher, S. J. (1988). Self-defeating behavior patterns among normal individuals: review and analysis of common self-destructive tendencies. *Psychological bulletin*, *104*(1), 3.

Beshears, J., & Milkman, K. L. (2011). Do sell-side stock analysts exhibit escalation of commitment? *Journal of Economic Behavior & Organization*, *77*(3), 304–317.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should i stay or should i go? how the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

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.

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical science*, *4*(3), 282–289.

Freeman, P. (1983). The secretary problem and its extensions: A review. *International Statistical Review/Revue Internationale de Statistique*, 189–206.

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *41*(2), 148–164.

Goldstein, D. G., McAfee, R. P., Suri, S., & Wright, J. R. (2020). Learning when to stop searching. *Management Science*, *66*(3), 1375–1394.

Halkjelsvik, T., & Rise, J. (2015). Persistence motives in irrational decisions to complete a boring task. *Personality and Social Psychology Bulletin*, *41*(1), 90–102.

Kelly, T. (2004). Sunk costs, rationality, and acting for the sake of the past. *Noûs*, *38*(1), 60–85.

Koole, S., & van't Spijker, M. (2000). Overcoming the planning fallacy through willpower: Effects of implementation intentions on actual and predicted task-completion times. *European Journal of Social Psychology*, *30*(6), 873–888.

Lai, T. L., Robbins, H., et al. (1985). Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, *6*(1), 4–22.

Lee, M. D. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive science*, *30*(3), 1–26.

McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological review*, *120*(2), 395.

Milyavskaya, M., & Werner, K. M. (2018). Goal pursuit: Current state of affairs and directions for future research. *Canadian Psychology/psychologie canadienne*, *59*(2), 163.

Mischel, W., Ayduk, O., & Mendoza-Denton, R. (2003). Sustaining delay of gratification over time: A hot-cool systems perspective.

Northcraft, G. B., & Neale, M. A. (1986). Opportunity costs and the framing of resource allocation decisions. *Organizational Behavior and Human Decision Processes*, *37*(3), 348–356.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives*, *23*(1), 6–16.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of Australian Mathematical Society*, *55*, 527–535.

Schelling, T. C. (1984). Self-command in practice, in policy, and in a theory of rational choice. *The American Economic Review*, *74*(2), 1–11.

Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "secretary problem". *Organizational behavior and human decision processes*, *69*(3), 221–236.

Steyvers, M., Lee, M. D., & Wagenmakers, E.-J. (2009). A bayesian analysis of human decision-making on bandit problems. *Journal of mathematical psychology*, *53*(3), 168–179.

Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald III, A. W., Thomas, M. J., & Redish, A. D. (2018). Sensitivity to "sunk costs" in mice, rats, and humans. *Science*, *361*(6398), 178–181.

Tenney, E. R., Logg, J. M., & Moore, D. A. (2015). (too) optimistic about optimism: The belief that optimism improves performance. *Journal of personality and social psychology*, *108*(3), 377.

Vermorel, J., & Mohri, M. (2005). Multi-armed bandit algorithms and empirical evaluation. In *European conference on machine learning* (pp. 437–448).

Wrosch, C., Scheier, M. F., Carver, C. S., & Schulz, R. (2003). The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial. *Self and identity*, *2*(1), 1–20.

Wrosch, C., Scheier, M. F., & Miller, G. E. (2013). Goal adjustment capacities, subjective well-being, and physical health. *Social and personality psychology compass*, *7*(12), 847–860.

Wrosch, C., Scheier, M. F., Miller, G. E., Schulz, R., & Carver, C. S. (2003). Adaptive self-regulation of unattainable goals: Goal disengagement, goal reengagement, and subjective well-being. *Personality and social psychology bulletin*, *29*(12), 1494–1508.

Zhang, Y., & Fishbach, A. (2010). Counteracting obstacles with optimistic predictions. *Journal of Experimental Psychology: General*, *139*(1), 16.

Appendix

Deriving the optimal strategy

In this section, we derive the optimal strategy for the pursue vs. quit task. We assume that the decision-makers' knowledge about the Bernoulli arms' success probability θ is given by a Beta distribution:

$$P(\theta|\alpha,\beta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}, \tag{1}$$

where α and β are hyperparameters encoding expectation about θ and the mean of the distribution is $\frac{\alpha}{\alpha+\beta}$. Further, we suppose that the decision-maker doesn't have any prior knowledge about the arms i.e., the hyperparameters of the prior (α_0,β_0) correspond to a flat prior, $(\alpha_0,\beta_0) = (1,1)$. We also considered scenarios with various different priors and find that our main predictions aren't significantly affected by the exact choice of the prior (changing α_0 and β_0 primarily shifts the threshold to be used when evaluating "good" options).

At each step t, the decision-maker takes an action a, transitions from state s_t to s_{t+1} and receives a reward r. The state s_t denotes the decision-maker's knowledge about the arm they are interacting with at time t. In the simplest case, s_t corresponds to the history of successes (i.e., outcomes where $r = 1$) and history of failures (i.e., outcomes where $r = 0$) of the arm. The decision-maker acquires and maintains a possibly probabilistic policy $\pi(s)$, which specifies a systematic mapping of states to distributions over actions. To compute the optimal policy, the decision-maker estimates the value of the current state s as:

$$V(s,T) = \mathrm{E}\left[\sum_{t=1}^{T} r_a(s_t) \mid s_t = s\right], \tag{2}$$

where T is the total number of pulls the decision-maker gets and s_t denotes the decision-maker's knowledge about the arm they are interacting with at time t. The value function can also be defined recursively as follows:

$$V(s,T) = \max_{a\in A}\left(r_a(s) + \sum_{s'\in S} p_a(s_t,s_{t+1})V(s_{t+1},T-1) \mid s_t = s\right), \tag{3}$$

where a is the decision to either persist with the current arm or to quit, A is the space of actions, and $r_a(s)$ is the expected reward for taking an action.

The space of states S consists of the decision-maker's knowledge about the current arm they are pulling, so it can be described with a pair (α,β) corresponding to a Beta distribution. The space of actions

Probability cap	0.9	0.8	0.7	Joint	More practice time
Success rate parameter	-0.49	-0.57	-0.62	-0.56	-0.66
Failure rate parameter	0.48	0.53	0.46	0.48	0.63
Remaining presses rate parameter	-0.008	-0.008	-0.004	-0.0061	-0.006
Logistic regression intercept	-1.3	-1.3	-1.4	-1.4	-1.6

Table 3

Logistic regression fits for different cap groups and the joint fit.

consists of either (a) pulling the current arm and moving to the state $(\alpha+1,\beta)$ with probability $\alpha/(\alpha+\beta)$ or moving to the state $(\alpha,\beta+1)$ with probability $\beta/(\alpha+\beta)$ or (b) quitting the current arm and moving to the next arm and pulling it thus arriving at one of the outcomes $(2,1)$ or $(1,2)$ with a probability $1/2$. Equation 3 can be thus re-written as:

$$V(s,T) = \max \begin{cases} \frac{\alpha}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}V(\alpha+1,\beta,T-1) + \frac{\beta}{\alpha+\beta}V(\alpha,\beta+1,T-1) \\ \frac{1}{2} + \frac{1}{2}V(2,1,T-1) + \frac{1}{2}V(1,2,T-1) \end{cases}. \tag{4}$$

The decision-maker can use Equation 4 to compute the optimal policy at each time step by simply comparing the first term to the second term – if the second term is greater than the first arm, then the decision-maker should quit the current arm and move to the next arm. Equation 4 can be evaluated numerically recursively for any T, which we do to derive the optimal strategy in the analysis presented in the main text.

Logistic regression fit in different probability cap groups and the longer practice round group

Table 3 compares fit parameters within different probability cap groups and the group with the higher-than-average practice round time. We can see that the rate parameters are close, and the fits have the same dependence on the number of successes, failures, and remaining presses.