Selective Sampling with Information-Storage Constraints*

Philippe Jehiel
PSE and UCL

Jakub Steiner
University of Zurich and CERGE-EI

August 20, 2018

Abstract

A decision-maker acquires payoff-relevant information until she reaches her storing capacity, at which point she either terminates the decision-making and chooses an action, or discards some information. By conditioning the probability of termination on the information collected, she controls the correlation between the payoff state and her terminal action. We provide an optimality condition for the emerging stochastic choice. The condition highlights the benefits of selective memory applied to the extracted signals. The constrained-optimal choice rule exhibits (i) confirmation bias, (ii) speed-accuracy complementarity, (iii) overweighting of rare events, and (iv) salience effect.

keywords: bounded rationality, cognitive constraints, information processing, stochastic choice, confirmation bias, speed-accuracy complementarity, probability weighting, salience.

JEL codes: D03, D80, D81, D83, D89, D90.

*This paper has been previously circulated under the title: "On Second Thoughts, Selective Memory, and Resulting Behavioral Biases." We thank Mark Dean, Andrew Ellis, Alessandro Pavan, Philip Reny, Larry Samuelson, Colin Stewart, Balazs Szentes, colleagues at the University of Edinburgh and the audiences at Bocconi, Queen Mary, Columbia, Ecole Polytechnique, Zurich and St Andrews universities, workshops and conferences in Erice, Alghero, Faro, Gerzensee, New York, Cambridge, Vancouver, and Barcelona. Ludmila Matysková and Jan Šedek provided excellent research assistance. Deborah Nováková and Laura Straková have helped with English. Jakub Steiner has received financial support from the Czech Science Foundation grant 16-00703S and from the ERC grant 770652. Philippe Jehiel thanks the ERC grant n° 742816 for funding.

1 Introduction

Economic agents often acquire information about the state of the economy before making their decisions. The information is typically modelled as a signal that helps the agent refine the distribution of the state and improve the decision-making. Often, signals come over time and agents can absorb only a small number of them. We capture this information-processing friction by assuming that agents can receive as many signals as they wish but can remember only a finite number of them when making their choices. In the simplest setting we analyze, the agent can only remember one signal. A key strategic variable that we consider is to allow the agent to ignore some signals with positive probability and restart the signal extraction process. An agent in the face of the first signal she observes can either make a choice based on this observation or dispose with the first signal and rerun the very same stochastic signal extraction process. In the face of the second observed signal she can either make a final decision (after the second observation and according to the same strategy that maps instantaneous observations to choices) or rerun the signal extraction again and so on. We allow agents to employ an arbitrary stationary decision process that specifies for each possible signal realization a probability with which the agent restarts the process as well as the chosen action in case of termination. We do not impose time constraints and costs in the basic formulation so that the friction comes solely from the limited information-storing capacity of the agent.

We ask ourselves: Should the agent optimally make her choice as soon as she receives the first signal whatever the realization of it is, or could she be better off by rerunning the very same information-acquisition process? Can hesitation—selective repetition of a fixed stochastic decision procedure when the agent cannot aggregate information—be welfare-enhancing?

A general insight is that selective rerunning of the primitive decision procedure is typically optimal. To document this most generally, we provide a simple necessary condition satisfied by the optimal rerunning strategy. The result is an interim indifference condition imposed on the agent who has concluded her decision-making with a plan to choose a particular action. Given the recommended action, the agent's posterior expected payoff from implementing this action must be the same as the posterior expected payoff from rerunning the whole decision-making—the whole selective repetitions of the primitive signal extraction—and implementing whichever action the second run of the decision-making will recommend. We refer to this as to the *second-thought-free* condition.

For illustration, consider a binary decision of whether to make an investment of a fixed size. The agent receives payoff 1 if she invests in the good state of the economy, payoff -2 if she invests in the bad state, and receives 0 when she does not invest whatever the state. Both states are a priori equally likely and give rise to a population of good and bad signals, with the share of the good signals at 90% in the good state and 10% in the bad state. The agent draws possibly several signal realizations in sequence but remembers only the last one when making her investment choice. As follows from simple optimization considerations, assume she invests if the last remembered signal was good (and does not invest when it is bad). Observe that the decision rule generated by the immediate termination upon the first signal that comes in does *not* generate a second-thought-free choice rule: An agent whose first observed signal was good prefers to rerun the decision process, since the new run will either lead to investing again or will give rise to the signal realization that conflicts with the first observation and will lead to not investing. Since, conditional on two conflicting signals, not investing is preferred (because the false-positive investment error is relatively costly), the agent benefits from having second thoughts when the first observed signal is good.

In the sequel, we interpret the probability of terminating the decision process after receiving a particular signal as a search intensity for the signal. A higher probability of termination at a given information set inflates the likelihood that the agent makes the terminal choice at the set. We show that the failure of the second-thought-free condition with uniform search intensity in the above investment decision example indicates that relative to the uniform search, the agent benefits from decreasing the search intensity for the good attribute.

More generally, the second-thought-free condition follows from the first-order condition imposed on the optimal search intensities. In the above example, a marginal increase in the relative search intensity in favor of the bad signal is welfare-enhancing when starting from the immediate termination strategy. Consider a deviation from the immediate termination strategy that consists of repeating the signal draw with a small probability whenever the observed signal is good. This new decision procedure effectively replaces a marginal measure of contingencies in which the agent terminates after observing the good signal with new draws of the signal. The indifference to marginal changes in the search intensities at optimum implies the second-thought-free condition. Given that a typical signal structure combined with the immediate termination strategy would not result in a second-thought-free rule for a generic payoff function, we conclude that some asymmetric termination strategy that differentiates the termination decision according to the last observed signal is generically optimal.

The model provides microfoundations to a range of behavioral stylized facts. The unifying principle of our behavioral insights is the intuition that the agent targets her search towards the type of evidence that would provide her with more informed posteriors under the uniform search. This principle generates *confirmation bias*, since evidence that confirms the agent's prior leads to more informed posterior than does evidence that contradicts the prior. An optimally targeted information search also generates *speed-accuracy complementarity*; that is, accuracy of choice declines with the response time. The effect is generated by the confirmation bias: The agent encountering evidence contradicting her prior is likely to disregard the evidence and to have a second thought. Hence, long response times indicate a surprising state of the world, and the constrained-optimal choice rule commits errors in the surprising state relatively often. *Overweighting of rare events* occurs when the agent's task is to form a probability belief about an event that is known to be rare, such as a flight accident, by observing a random flight outcome. Since observing a flight accident is far more informative about the probability of future accidents than observing an uneventful flight, the agent optimally biases her search towards eventful flights. In the last behavioral application, we show that distinct states of the world are *salient* in the sense that they attract the agent's attention (i.e., trigger higher termination rates in our framework). The effect arises because an indistinct perception stimulus that can be generated by several similar states is less informative than a distinct perception stimulus likely generated by a specific distinct state. Hence, the optimal information search targets stimuli indicating distinct states.

While our leading interpretation is in terms of a single-person decision process with information storage limitations, we note that there are alternative organizational interpretations of our model. We offer one such alternative interpretation in Section 5.1.2 where we briefly discuss the confirmation bias arising under the optimal editorial policy of media outlets that maximize welfare of its readers who have limited attention span. Another interpretation is that of a firm in which a sequence of employees acting one after the other can acquire private information, make a decision, or delegate the decision to the next employee without the ability to pass the information to the latter.

When the decision-maker can control error distributions, the stochastic process that governs the pattern of choices is influenced by the decision-maker's payoffs; that is, it is shaped by her incentives. At various levels of formalization, such an idea of error- management has appeared repeatedly in biology, psychology and economics; see Johnson et al. (2013) for an interdisciplinary review. A range of economic models formalize the error-management idea by specifying particular sets of feasible error distributions. These models differ greatly in the frictions imposed on decision-making, and hence in the predicted stochastic choice rules: Rational inattention models, e.g. Sims (2003), Matějka and McKay (2015) and Steiner et al. (2017), impose an abstract, entropy-based friction on information processing. A distinct information-processing friction is assumed in sequential-sampling and drift-diffusion models that explicitly study continuous evolution of beliefs, e.g. Wald (1945), Arrow et al. (1949), Ratcliff (1978), Hébert and Woodford (2016), and Morris and Strack (2017), whereas Che and Mierendorff (2016) and Zhong (2017) study discontinuous learning counterparts to those models. Yet another modeling approach to limited cognition is based on finite automata, e.g. Hellmann and Cover (1970), Compte and Postlewaite (2012), Wilson (2014), and Basu and Chatterjee (2015). Compared to the above diverse models, our approach focuses on a simpler instrument: the termination probability at each information set. This allows us to derive simple necessary conditions that apply to all models in which such an instrument is available, which turns out to be all that is needed to derive connections to psychological biases as reported above.

This paper belongs to a growing economic literature that explains behavioral stylized facts as the constrained-optimal behavior of decision-makers facing information processing frictions. For instance, Robson (2001), Rayo and Becker (2007), Netzer (2009), and Khaw et al. (2017) provide microfoundations for risk attitudes; Gabaix and Laibson (2017) endogenize discounting; and Wilson (2014), Compte and Postlewaite (2012), and Leung (2017) establish constrained-optimal ignorance of weakly informative news. The partitional model of Dow (1991) focuses on the problem of coarsening of rich information available to the decision-maker.[1]

[1]Somewhat less related is a literature that explores how exogenous errors in learning lead to behavioral biases; see coarse learning in Jehiel (2005) and its application to overoptimism in Jehiel (2018). By contrast, in our approach,

Our model can be related to the literature on decision making with forgetful decision makers as considered in Piccione and Rubinstein (1997). Our approach corresponds to their ex ante approach, and our derivation of the second-thought free condition that applies to it can be related to the observation in their absent driver paradox leading example that the ex ante optimal solution is also a (modified) multi-self equilibrium in which the decision problem is viewed as a team between multiple selves all sharing the decision-maker's objective where each self maximizes the interim expected utility of the decision maker internalizing the other selves' strategies in her Bayesian inference.

Meyer (1991) studies optimal biases in a sequential-learning problem of an agent who receives a sequence signals and, unlike our agent, can aggregate the sequence. Meyer's main insight is that some asymmetries in the signal structure are optimal. Although optimal asymmetries arise both in her and our frameworks, the two papers study distinct optimizations. While our agent controls termination probabilities in a stationary decision process, Meyer's agent controls the choice of a Blackwell experiment in each round of a non-stationary process.

A related model in which the decision-maker over-samples certain type of evidence was developed by psychologists Lieder et al. (2014). This project and us examine frictions on the opposing sides of information processing. While we study non-representative sampling during the information gathering and allow for frictionless formation of the posterior beliefs, Lieder et al. take available information as given and impose the friction on expectation formation. Their decision-maker forms the expected payoffs in a Monte Carlo simulation as an average over a finite sample of simulated payoff states. She optimally oversamples states in which stakes are high.

2 Model

2.1 General model

An agent faces a decision under uncertainty. She chooses an action $a \in A$ in a process specified below and receives a payoff $u(a, \theta)$ in the fixed payoff state $\theta \in \Theta$ drawn from an interior prior distribution $\pi \in \Delta(\Theta)$. The state space Θ and the action set A are finite. The agent chooses a signal structure—or equivalently a Blackwell experiment—p, where p is characterized by a family of conditional signal distributions $p(x \mid \theta)$, $\theta \in \Theta$. The experiment generates a signal realization x from a finite signal space X. The conditional signal distributions are fully mixed: $p(x \mid \theta) > 0$ for all x, θ. We allow the agent to choose among possibly several such Blackwell experiments and we let $\mathcal{P}$ denote the set of experiments from which she chooses. We impose no restrictions on the set $\mathcal{P}$ of the feasible experiments (other than the full-support of each p). One interpretation is that each experiment $p \in \mathcal{P}$ is a particular reasoning approach available to the agent.

The agent can repeat the selected experiment arbitrarily many times, but she is unable to aggregate the information across the repetitions. Each run of the experiment is a cognition that exhausts the agent's capacity dedicated to the problem being solved. Once the agent hits the

the agent optimizes the error distribution given the constraints.

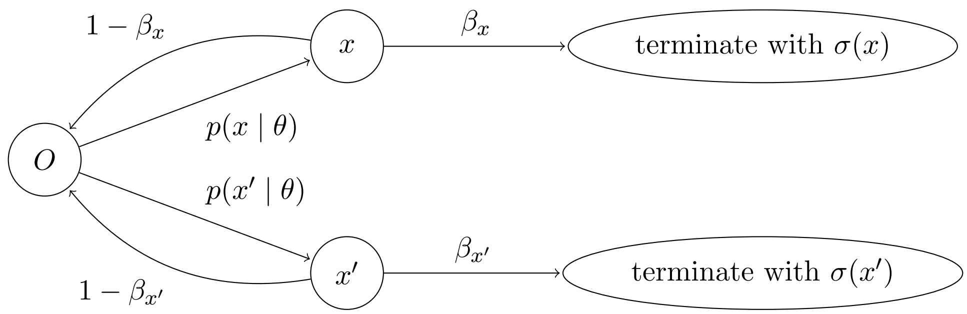

Figure 1: For each (p, β, σ), the decision process is a Markov chain evolving on the agent's states of mind, with transition probabilities that depend on the payoff state θ. The chain begins in the state of mind O and transits to states $x \in X$ with probabilities $p(x \mid \theta)$. The process returns to O with probability $1 - \beta_x$, or terminates with choice of $a = \sigma(x)$ with probability β_x.

constraint at the end of the experiment, she can continue only after she unclogs her capacity by amnesia.

We model this as follows. The agent can condition the repetition of the experiment on the last observed signal realization. She chooses a vector $\beta = (\beta_x)_{x \in X} \in B = [0,1]^{|X|} \setminus \{(0, \ldots, 0)\}$ of termination probabilities β_x for each signal realization x; we call β a *termination strategy*. The agent runs the experiment p for the first time, receives signal realization x_1 with probability $p(x_1 \mid \theta)$ and terminates the reasoning with probability β_{x_1}. She restarts her reasoning with the complementary probability $1 - \beta_{x_1}$, and receives a signal realization x_2 from a new run of the process p with probability $p(x_2 \mid \theta)$, terminates with probability β_{x_2} or restarts with probability $1 - \beta_{x_2}$, and continues to rerun the primitive process p until she terminates after a random number of repetitions of p; see Figure 1. When the agent chooses having distinct β_x for different x, then she implements the familiar idea of selective memory; some facts and observations are easily forgotten whereas others are remembered and they trigger choice. After the agent terminates the reasoning with a terminal signal realization x, she selects an action $a = \sigma(x)$ according to an *action strategy* $\sigma : X \longrightarrow A$.[2] Let S be the set of all mappings from X to A.

By excluding the termination strategy $(0, \ldots, 0)$ we prohibit the agent from avoiding to take the decision $a \in A$. Since $\beta \neq (0, \ldots, 0)$ and each feasible experiment p generates all signal values with a positive probability in each state, the decision process almost surely eventually terminates. Note that we can accommodate outside option by enlarging the action space.

The outcomes of distinct runs of p are conditionally independent. Thus, the probability that the agent terminates after t repetitions of the experiment p resulting in the signal history $\mathbf{x}^t = (x_1, \ldots, x_t)$ is

$$\rho\left(\mathbf{x}^t \mid \theta; p, \beta\right) = \beta_{x_t} p(x_t \mid \theta) \prod_{l=0}^{t-1} \left(1 - \beta_{x_l}\right) p\left(x_l \mid \theta\right). \tag{1}$$

[2] We do not allow for mixed action strategies since the optimum can always be achieved with a pure action strategy.

We let

$$r(a \mid \theta; p, \beta, \sigma) = \sum_t \sum_{\mathbf{x}^t : \sigma(x_t) = a} \rho\left(\mathbf{x}^t \mid \theta; p, \beta\right) \tag{2}$$

denote the probability that the agent who employs the experiment p, the termination strategy β, and action strategy σ terminates with action a. We call $r(p, \beta, \sigma) := (r(a \mid \theta; p, \beta, \sigma))_{a \in A, \theta \in \Theta}$ the *choice rule*. The set of feasible choice rules is $\mathcal{R}(\mathcal{P}) = \{r(p, \beta, \sigma) : p \in \mathcal{P}, \beta \in B, \sigma \in S\}$. Sometimes we abuse notation, omit p, β, σ and write $r(a \mid \theta)$ for the probability of a in state θ under the rule constructed by some p, β, σ.

The *repeated-cognition problem* is to select a feasible choice rule r that maximizes the expected payoff:

$$\max_{r \in \mathcal{R}(\mathcal{P})} \sum_{\theta \in \Theta, a \in A} \pi_\theta r(a \mid \theta) u(a, \theta). \tag{3}$$

The agent solving the repeated-cognition problem knows the prior π, payoff function u, and the set $\mathcal{P}$ of the feasible processes p. The optimization in (3) can be an outcome of selective pressures that favor successful decision procedures via cultural or biological evolution, or via competition of firms differing in their internal procedures.

We focus on the benefit of the repeated cognition in that our baseline model abstracts from the cost of time and is therefore applicable to agents who can repeat the basic cognition process p quickly. The model extends to agents who exponentially discount future payoffs, and thus face non-trivial cost of repeated cognition. We report such an extension in Section 6.2.

2.2 Examples

To demonstrate the flexibility of the general model, we now informally discuss three specific settings. The first one, the elementary setting, is a highly stylized setting in which the agent can repeatedly employ only a single statistical experiment and can remember only one signal realization. We use this setting for pedagogical purposes in the applications in Section 5. The second and third settings feature sophisticated agents with non-trivial memory that she can use to aggregate information over many observed signal realizations. Perhaps surprisingly, we show in Section 6 that those settings with non-trivial memory can in fact be interpreted as special cases of our general model that on its face value allows only for trivial memory. We show that such accommodation of non-trivial memory is possible via expansion of the set $\mathcal{P}$ of the primitive experiments.

Moreover, when the state and action spaces are binary, then all the three settings have a same analytical solution and the optimal stochastic choice rule in the two sophisticated settings differs from the solution of the elementary setting only in a value of a one-dimensional parameter that summarizes the agent's ability to process information.

Example 1 (*elementary setting*). The agent has access only to one statistical experiment: $\mathcal{P} = \{p\}$, where p is an exogenous information structure that specifies conditional signal distributions $p(x \mid \theta)$. In this case, the agent only chooses the termination probabilities β_x, and the action strategy σ that determines the action $a = \sigma(x)$ chosen at each terminal signal realization x.

The formalization and accommodation of the next two examples within our baseline model is nontrivial and postponed to Section 6.1.

Example 2 (*imperfect information aggregation*). This setting relaxes the agent's inability to aggregate information across the repetitions of her reasoning by endowing her with a finite set of memory states that she can use to represent the signal histories. The setting of this example builds on Hellmann and Cover (1970) and Wilson (2014). The agent can repeatedly sample from a single statistical experiment that generates signal realizations from a finite signal space. Additionally, the agent is endowed with a finite set of memory states. After each run of the experiment, the agent randomizes between terminating and continuation of the decision process, where in the latter case, she may transition to a new memory state. The termination decisions and the transitions among memory states follow a stationary mixed strategy that conditions on the current memory state and the last observed signal. Once the agent terminates, she maps the last memory state and the last observed signal value to a chosen action. The feasible statistical experiment and the set of the memory states specify a set of constructible choice rules, from which the agent chooses the one that maximizes her ex ante expected payoff.

Example 3 (*partial forgetting*). The agent of this example can remember up to a fixed finite number of signal realizations generated by a single statistical experiment. In each round of her decision process, she can discard a subset of the currently remembered signals values, extract a new signal realization, or terminate, where each of these decisions is determined by a stationary mixed strategy that conditions on the currently remembered stock of the signal values. The statistical experiment and the maximal number of the signals that the agent can remember determine the set of the stochastic choice rules that she can construct, from which she chooses the rule that maximizes her ex ante expected payoff.

3 Optimal cognition biases

We now derive a necessary optimality condition that the choice rule solving the repeated-cognition problem must satisfy. We argue in Section 3.3 that, generically, the condition requires the agent to engage in selective information processing—that is, to ignore some signals more often than others.

3.1 Second-thought-free choice rules

We start with a definition of *second-thought-free* choice rules. If the agent's decision process generates such a rule, then she has no incentive to rerun the process regardless of the action recommendation with which the process terminates. Our main result below states that an optimal rule that solves the repeated-cognition problem is second-thought-free.

Let A and Θ be finite action and state sets with generic elements a and θ. Let r be a generic stochastic choice rule that specifies conditional probabilities $r(a \mid \theta)$ of each action $a \in A$ in each state $\theta \in \Theta$.

Definition 1. *The choice rule r is* second-thought-free *with respect to the utility u and prior π if the agent prefers each action recommended by the rule to a new run of the rule r. That is, for each action a chosen with positive probability,*

$$\mathrm{E}_\alpha[u(a_1, \theta) \mid a_1 = a] \geq \mathrm{E}_\alpha[u(a_2, \theta) \mid a_1 = a], \tag{4}$$

where the expectations are with respect to the random variables θ and a_2, and $\alpha(\theta, a_1, a_2) = \pi_\theta r(a_1 \mid \theta) r(a_2 \mid \theta)$ is the joint distribution of the state and two actions consecutively generated by r.

The definition requires the agent who terminates her decision process with an action plan a to weakly prefer the action a to forgetting a and choosing whichever action a new run of the decision process will recommend. Although the definition allows for the strict preference against having a second thought, the next lemma shows that (4) is always met with equality: If a choice rule is second-thought-free, then the agent is indifferent between terminating and the second thought. The lemma is a simple consequence of the law of iterated expectations.

Lemma 1. *If a choice rule r is second-thought-free, then the condition* (4) *is met with equality for each action a chosen with positive probability:*

$$\mathrm{E}_\alpha[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}_\alpha[u(a_2, \theta) \mid a_1 = a]. \tag{5}$$

All proofs are relegated to Appendix. We refer to (5) as the *second-thought-free condition.*

3.2 Optimality condition

We provide here a general necessary optimality condition imposed on the stochastic choice rule. Note that the existence of the solution to the repeated-cognition problem is not guaranteed since we do not impose any restrictions on the set $\mathcal{P}$ of the feasible experiments. We prove existence for a binary setting in Section 4, where we impose more structure on the model.

Proposition 1. *If a choice rule solves the repeated-cognition problem* (3)*, then it is second-thought-free and satisfies* (5)*.*

To understand the statement, consider the optimal choice rule r^* generated by a process that consists of a random number of repetitions of a primitive cognition p. Once these repetitions of p terminate with a terminal signal realization x and the agent is about to take an action $a = \sigma(x)$, then, according to the proposition, she must be indifferent between a, and running the process associated with r^* from scratch, where the new run of r^* would involve new repetitions of p.

To prove Proposition 1, we first introduce an *effective experiment* $s(p, \beta)$ and distinguish it from the *primitive experiment* p. While $p(x \mid \theta)$ specifies the probability that one run of the experiment p results in signal realization x, $s(x \mid \theta; p, \beta)$ is the probability that selective repetitions of p according to the termination strategy β terminate with x. Relative to the primitive probabilities $p(x \mid \theta)$, the effective probabilities $s(x \mid \theta; p, \beta)$ are inflated for those x at which the agent terminates with a high probability β_x:

Lemma 2. *The probability that the agent who employs a primitive experiment* p *and a termination strategy* β *terminates with* x *in state* θ *is*

$$s(x \mid \theta; p, \beta) = \frac{\beta_x p(x \mid \theta)}{\sum_{x' \in X} \beta_{x'} p(x' \mid \theta)}. \tag{6}$$

The simple proof in Appendix exploits the stationarity of the decision process. The lemma implies that $s(p, \beta)$ and hence also $r(p, \beta, \sigma)$ are homogeneous of degree zero with respect to β. Thus, since we abstract from the delay costs, for any optimal termination strategy β^*, $\alpha\beta^*$ for $\alpha \in (0, 1)$ is optimal too, and it generates the same optimal choice rule r^* as β^*. This multiplicity of implementation of the optimal choice rule would disappear in a natural approximation of our model with exponential discount factor approaching 1. Such approximation would select the quickest available decision process that implements the optimal feasible rule r^*; that is, it would impose that $\max_{x \in X} \beta_x = 1$.

Lemma 2 can be used to rewrite the agent's objective as an explicit function of the termination strategy: The repeated-cognition problem is equivalent to

$$\max_{p \in \mathcal{P}, \beta \in B, \sigma \in S} \sum_{\theta \in \Theta, x \in X} \pi_\theta \frac{\beta_x p(x \mid \theta)}{\sum_{x' \in X} \beta_{x'} p(x' \mid \theta)} u(\sigma(x), \theta). \tag{7}$$

Proposition 1 follows from the first-order condition with respect to the termination strategy β. Since optimization over the experiments and over the action strategies plays no role in the result, Proposition 1 continues to hold if p and σ are suboptimal, as when the action strategy and employed experiment are hardwired automatic responses to environment.

Comment. Proposition 1 does not make reference to the set $\mathcal{P}$ of the feasible experiments. Rather, the second-thought-free condition is imposed only on the choice rule r^*, on the prior, and on the utility function. Thus, assuming the analyst has access to rich data (about the state and the action) generated by the decision process, he can check if the second-thought-free condition is satisfied even if he is ignorant of the set $\mathcal{P}$ to the extent he knows the utility function. And if the analyst does not know the utility function, he can refine his estimate of it, assuming the second-thought-free condition is satisfied.

3.3 Benefit of cognition biases

We now illustrate in an example how the agent benefits from repeated signal extraction in absence of information aggregation. The key to such beneficial "hesitation" process is a cognition bias that favors termination of the decision-making after encountering a certain type of signal realization.

The agent's task is to announce the realized value of the binary state $\theta \in \{0, 1\}$. She is rewarded with payoff $u(a, \theta) = 1$ if her announcement $a \in \{0, 1\}$ matches the state θ, and she receives zero otherwise. The state θ is drawn from the uniform prior and the agent has access only to one statistical experiment; $\mathcal{P} = \{p\}$ is a singleton. The available statistical experiment is asymmetric in that it satisfies[3]

$$.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .6.$$

That is, $x = 1$ (resp. $x = 0$) is informative that the state θ is more likely to be 1 (resp. 0) and the signal realization 1 is more informative than realization 0. Let us fix the action strategy to be the identity function $\sigma_I(x) = x$.[4] With probability $(.5 - .4)/(.9 - .4) = .2$, the experiment delivers the signal value 1, which provides the agent with the correct advice in 90% of cases, whereas the somewhat less informative signal realization 0 that is correct in only 60% of the cases is observed with probability .8.

Consider first an agent who always terminates immediately after the first observed signal realization: $(\beta_0, \beta_1) = (1, 1) = \mathbf{1}$. For this agent, the choice rule $r(p, \mathbf{1}, \sigma_I)$ and the primitive experiment p coincide, and the agent makes the correct choice with probability $.2 \times .9 + .8 \times .6 = .66$. The choice rule $r(p, \mathbf{1}, \sigma_I) = p$, however, is not optimal since it fails to be second-thought-free. To verify this, consider the agent who has run the experiment p, receives the relatively unreliable signal realization 0, and is about to conclude the decision-making with action $a_1 = 0$. Before she implements a_1, let her contemplate having a second thought that involves running the experiment p once again and implementing whichever action $a_2 = x_2$ the new run recommends. If $a_2 = a_1$, then the second thought will have been inconsequential. If $a_2 \neq a_1$, then it is more likely that $\theta = 1$ than $\theta = 0$ (because $x = 1$ is more informative than $x = 0$), and thus the agent benefits from switching her choice from $a_1 = 0$ to $a_2 = 1$. Thus, overall, the agent who is about to terminate the decision process $r(p, \mathbf{1}, \sigma_I)$ with action 0 benefits from having the second thought. By Proposition 1, the process $r(p, \mathbf{1}, \sigma_I)$ is not optimal.

Generically, choice rules induced by immediate termination are not second-thought-free, and therefore, like in this example, the agent who has access to a generic experiment p and has a generic utility u and a prior π benefits from selective repetitions of the experiment p. In this example, due to the symmetry of the prior and of the utility function, the second-thought-free rule must provide correct action recommendation with a same probability in both states: the optimal termination strategy β^* must satisfy symmetry:

$$r(1 \mid 1; p, \beta^*, \sigma_I) = \frac{\beta_1^* p(1 \mid 1)}{\beta_0^* p(0 \mid 1) + \beta_1^* p(1 \mid 1)} = \frac{\beta_0^* p(0 \mid 0)}{\beta_0^* p(0 \mid 0) + \beta_1^* p(1 \mid 0)} = r(0 \mid 0; p, \beta^*, \sigma_I). \tag{8}$$

To see why the optimal rule must satisfy such symmetry, recall that the second thought affects the payoff only if the two runs of the decision process would result in the conflicting signal realizations $x_1 \neq x_2$. If the rule was not symmetric, then the agent would fail to be indifferent between the two actions in this contingency, and hence would want to have a second thought after terminating with one of the two actions. Vice versa, when the rule satisfies the symmetry (8), then both states are equally likely in this contingency, and thus the agent is indeed indifferent between the two actions; the symmetric rule is second-thought-free.

[3]The specified posteriors and the prior pin down the experiment p: $p(1 \mid 1) = .36$ and $p(0 \mid 0) = .96$.

[4]This strategy is shown to be optimal in Section 4.

The condition (8) implies that $\beta_0^*/\beta_1^* = 0.15$. The fastest optimal rule terminates as soon as the agent observes the highly informative signal realization 1, but when the agent observes the unreliable realization 0, then she is likely to hesitate: she reruns her cognition p with probability $1 - .15 = .85$. The optimal choice rule is $r(1 \mid 1; p, \beta^*, \sigma_I) = r(0 \mid 0; p, \beta^*, \sigma_I) = 0.79$ and thus the overall expected payoff increases to 0.79 compared to only .66 under the decision process that terminates immediately.

4 Analytical solution of the binary setting

We extend the example studied in subsection 3.3 to cover all settings with binary action and state spaces. The agent chooses $a \in A = \{0, 1\}$ and receives $u(a, \theta)$, where the payoff state θ is drawn from an interior prior $\pi \in \Delta(\Theta)$, and $\Theta = A$. To avoid a trivial case, we assume that neither action is dominant. Then, without loss of generality, $u(a, \theta) = u_\theta > 0$ if $a = \theta$ and $u(a, \theta) = 0$ otherwise. The set $\mathcal{P}$ of the feasible statistical experiments is finite, and each $p \in \mathcal{P}$ delivers a signal realization x from a finite signal space X with probability $p(x \mid \theta) > 0$. The agent chooses the statistical experiment, the termination strategy and the action strategy to maximize her expected payoff.

We introduce a summary statistics $\overline{d}$ of the information structure $\mathcal{P}$,

$$\overline{d} = \max_{p \in \mathcal{P}, x \neq x' \in X} \frac{p(x \mid 0) p(x' \mid 1)}{p(x \mid 1) p(x' \mid 0)}, \tag{9}$$

and note that $\overline{d} \geq 1$. When $\overline{d}$ is large then there exists an experiment p and two signal realizations x and x' such that one of the realizations is relatively strongly indicative of the state 1 and the other realization of the state 0. We refer to $\overline{d}$ as to the perceptual distance between the two states. It turns out that for the purpose of the analytical characterization of the optimal choice rule, $\overline{d}$—a one-dimensional real number—is a sufficient summary statistics for the complex, multidimensional information structure $\mathcal{P}$. This summary statistics is independent of the agent's incentives and of the prior.

We assume for simplicity that the maximizer in (9) is unique, and note that the uniqueness is generic.[5] We denote the maximizing experiment by $\overline{p}$, label the maximizing pair of the signal values by 0 and 1, and without loss of generality choose the signals labels in such the way that the likelihood ratio $p(1 \mid \theta)/p(0 \mid \theta)$ increases in θ. Signal 1 is thus relatively indicative of the state 1. Finally, we introduce parameter $R = \frac{\pi_1 u_1}{\pi_0 u_0}$ that summarizes the incentives and prior; it measures the relative a priori attractiveness of action 1.

Proposition 2. *1. When $R \geq \overline{d}$, then the agent always chooses action 1;*

2. when $R \leq 1/\overline{d}$, then the agent always chooses action 0;

[5] When there are multiple maximizers (p, x, x') then each of them induces a solution to the repeated cognition problem. All these solutions coincide in the optimal stochastic choice rule $r^*(a \mid \theta)$.

3. when $R \in (1/\overline{d}, \overline{d})$, then the agent chooses both actions with positive probabilities and

$$r^*(1 \mid 1) = \frac{\overline{d}R - \sqrt{\overline{d}R}}{(\overline{d}-1)R},\ r^*(0 \mid 0) = \frac{\overline{d} - \sqrt{\overline{d}R}}{\overline{d}-1}, \tag{10}$$

$$\frac{\beta_1^*}{\beta_0^*} = \frac{\overline{d}R - \sqrt{\overline{d}R}}{\sqrt{\overline{d}R} - R} \frac{\overline{p}(0 \mid 1)}{\overline{p}(1 \mid 1)}. \tag{11}$$

Thus, for the binary setting, the solution exists, and the optimal stochastic choice rule is unique. The ratio β_1^*/β_0^* of the termination probabilities is unique too, if both actions are chosen with positive probabilities. When the ex ante attractiveness of one of the actions is too strong relative to the perceptual distance of the two states, then the agent always chooses the ex ante attractive action. The decision process is non-trivial for intermediate incentives: the agent engages in repeated cognition and she chooses both actions with positive probabilities. In Section 6.1 we extend Proposition 2 to cover the agents from Examples 2 and 3 endowed with elaborate memory.

The solution of the repeated-cognition problem has natural comparative statics reminiscent of wishful thinking. As the ex ante attractiveness of a state increases, the agent is more willing to accept evidence in support of this state and ignore evidence in support of the alternative state, and the relative speed of the decision-making increases with the attractiveness of the state. Let $f_\theta = \sum_x \beta_x p(x \mid \theta)$ stand for the decision rate in state θ; it is the per-round probability that the agent in state θ terminates the decision-making.

Corollary 1. *Let $R \in (1/\overline{d}, \overline{d})$. Then,*

1. *relative willingness β_1^*/β_0^* to terminate with the signal value 1 increases with R;*

2. *relative decision rate f_1/f_0 increases with R.*

Let us sketch the main ideas from the proof of Proposition 2. Consider first an agent who can employ a single binary statistical experiment $p(x \mid \theta)$, $X = \{0, 1\}$, and who uses action strategy $\sigma_I(x) = x$. All choice rules r that this agent can construct satisfy a simple feasibility condition

$$\frac{r(1 \mid 1) r(0 \mid 0)}{r(1 \mid 0) r(0 \mid 1)} = \frac{\frac{\beta_1 p(1|1)}{\sum_x \beta_x p(x|1)} \frac{\beta_0 p(0|0)}{\sum_x \beta_x p(x|0)}}{\frac{\beta_1 p(1|0)}{\sum_x \beta_x p(x|0)} \frac{\beta_0 p(0|1)}{\sum_x \beta_x p(x|1)}} = \frac{p(1 \mid 1) p(0 \mid 0)}{p(1 \mid 0) p(0 \mid 1)} := d_p.$$

For this agent, the solution for the optimal rule is implied by the above feasibility condition and the second-thought-free condition. We show in the proof that when the feasible experiments attain many signal values, then the agent terminates at two most informative signal values and effectively ignores all other signal realisations x by setting $\beta_x = 0$. Then, the solution for the general information structure $\mathcal{P}$ equals the solution for the binary experiment with the highest value $d_p = \overline{d}$ that the agent can achieve by selecting $p \in \mathcal{P}$ and a pair of signal values x, x' at which she terminates with positive probabilities.

In the next section, we apply the analytical solution derived in Proposition 2 to several behavioural phenomena.

5 Behavioral applications

This section presents four behavioral effects generated by our model. We demonstrate the first three effects—confirmation bias, speed-accuracy complementarity, and overweighting of rare events—in the binary setting from Section 4. The fourth effect—salience of distinct states—will be presented in a setting with multiple states.

5.1 Confirmation bias

5.1.1 Confirmation bias in individual information-processing

Psychologists and economists distinguish at least three mechanisms leading to confirmation bias: (i) People search for evidence selectively, targeting the evidence type in accord with their priors, e.g. Nickerson (1998); (ii) they selectively memorize and recall the data supporting their priors, e.g. Oswald and Grosjean (2004); and (iii) they selectively interpret ambiguous evidence, e.g. Rabin and Schrag (1999) and Fryer et al. (2016). We focus on the first two mechanisms and interpret them in light of our optimal repeated-cognition result. We find our model well suited for the study of the confirmation bias since it focuses on agent's choice whether to ignore or process the encountered evidence.

The next result provides a simple illustration of why some form of confirmation bias is constrained optimal. We consider here an agent who has access to only one symmetric primitive experiment.

Corollary 2. *When action 1 is a priori more attractive, $R \in (1, \overline{d})$, and the unique primitive experiment is symmetric, $p(1 \mid 1) = p(0 \mid 0)$, then the agent searches relatively more intensively for signal value 1: $\beta_1^* > \beta_0^*$.*

To see how the above result relates to confirmation bias, consider an agent whose task is to announce the realized state of the world: she receives reward $u_1 = u_0 = 1$ if she makes the correct announcement and receives 0 otherwise. The agent finds the state $\theta = 1$ a priori more likely than the state 0, $\pi_1 > \pi_0$. Consider in this setting the suboptimal decision process that terminates immediately after the first run of the experiment and chooses the action equal to the observed signal value: $\beta_0 = \beta_1 = 1$, $\sigma = \sigma_I$. We first observe, paralleling an argument made in subsection 3.3, that such an unbiased process is suboptimal. To see this, assume that the agent has observed the a priori unlikely signal value 0. Such a surprised agent is better off by restarting the decision-making instead of terminating with action 0, since if the new run of the process concludes with signal value and action 1, then the switch from action 0 to 1 is beneficial. This is because when the experiment p is symmetric, then, conditional on the two conflicting signal values, the a priori more common state 1 is relatively more likely. The agent benefits from the second thought whenever she receives the surprising recommendation, and thus will deviate from the uniform search in favor of the a priori likely signal value 1.

The optimal strategy resembles the natural process in which the selective memory gives rise to confirmation bias. We consider the fastest optimal strategy, letting $\beta_1^* = 1$. When the agent observes signal value 1 that confirms her prior belief, then she terminates and immediately announces the state 1. But if she is surprised, observing signal value 0 that contradicts her prior, then she discards the signal with positive probability β_0^* and repeats the experiment. Although finding the exact optimal value β_0^* may be difficult, the fact that double-checking one's own reasoning when one arrives at a surprising conclusion is a common practice suggests that people are able to deviate from the unbiased information-acquisition process in the payoff improving direction.

5.1.2 Confirmation bias in media

Let us relate the above effect of confirmation bias to the choice of media outlets. For this application, we have in mind a multi-person interpretation of the model. Let each state of the world $\theta \in \{0, 1\}$ generate an infinite sequence of signal values x_k iid. drawn from the conditional distribution $p(x \mid \theta)$. The agent, whose task is to identify the realized state, can comprehend only one such signal realization and she can access it only via a media outlet. The media outlets differ in their editorial policies $(\beta_x)_x$. Each outlet draws a first signal realization x_1, terminates and reports x_1 to its readers with probability β_{x_1}, and with the residual probability $1 - \beta_{x_1}$ the outlet redraws the new signal realization x_2, etc, until the outlet terminates its search and reports the last observed signal value. The reader of the outlet with an editorial policy β observes signal value x in the state θ with probability $s(x \mid \theta; p, \beta) = \frac{\beta_x p(x|\theta)}{\sum_{x'} \beta_{x'} p(x'|\theta)}$.

As in Gentzkow and Shapiro (2006), our agent prefers outlets biased in favor of her prior belief; she prefers an editorial search policy β that is biased in favor of the signal realization that confirms the agent's own prior. For example, a voter a priori favouring, say, Trump, who has time to read only one headline in the media outlet of her choice, will optimally choose an outlet that persistently searches for evidence favorable to Trump, and that will report encountered evidence that is unfavorable to Trump only with a relatively smaller probability.

The source of the media bias in Gentzkow and Shapiro is reputation: Their agents evaluate media outlets confirming their prior beliefs as being relatively reliable information sources. In our case, all outlets are ex ante identical in that they have access to the same signal-generating process and thus reputation does not play a role. Rather, in our case, the demand for the prior-confirming outlets is driven by the information-aggregation friction. When the reader's attention span allows the outlets to report only one signal value, then the optimal editorial policy favors the signal value that advises the reader correctly in the a priori more likely state, since this state has a large weight in the reader's a priori objective.[6]

[6] See also Calvert (1985), Suen (2004), and Che and Mierendorff (2016) for constrained-optimal media-bias models.

5.2 Speed-accuracy complementarity

Our model generates the speed-accuracy complementarity effect—a stylized fact stating that delayed choices tend to be less accurate than speedy choices; see the psychology studies of Swensson (1972) and Luce (1986). We establish this effect in the setting from the previous subsection: $u_0 = u_1 = 1$, $\pi_1 > \pi_0$, considering again a symmetric primitive signal distribution, $p(1 \mid 1) = p(0 \mid 0) > 1/2$.

Let $\varphi(\theta, a, t)$ be the joint probability distribution of the state θ, chosen action a, and the reaction time t generated by the solution (p, β^*, σ_I) of the repeated-cognition problem.

Corollary 3. *The probability* $\Pr_\varphi(a = \theta \mid t)$ *of the correct choice decreases with response time* t*.*

Due to the stationarity of the decision process, the probability of the correct choice conditional on the payoff state is independent of the reaction time: $\Pr_\varphi(a = \theta \mid \theta, t) = \Pr_\varphi(a = \theta \mid \theta)$. At optimum, this conditional probability of the correct choice is larger in the a priori more likely state 1 than in the state 0, reflecting the relative weights of the two states in the a priori objective. Overall, unconditionally on the payoff state, the probability $\Pr_\varphi(a = \theta \mid t)$ of the correct choice depends on the response time because t correlates with θ. A long response time indicates that the agent has repeatedly encountered the a priori surprising signal value 0 and has hesitated to terminate. Hence, conditional on large t, the likelihood of the a priori surprising state becomes high. The longer the agent has hesitated, the more likely it is that she is facing the a priori surprising state in which she is making more mistakes.

5.3 Overweighting of rare events

We continue to consider the state-recognition task; $u_0 = u_1 = 1$. In contrast to previous applications, we assume that the two states $\theta = 0, 1$ are equally likely, but the distribution of the signal values $x = 0, 1$ is asymmetric across states. Specifically, the probability of $x = 1$ in the state θ is $\rho_\theta \in (0, 1)$ and the probability of $x = 0$ is $1 - \rho_\theta$. We assume, essentially without loss of generality, that $\rho_0 < \rho_1 < 1 - \rho_0$.[7] The a priori probability of event $x = 1$ is $(\rho_0 + \rho_1)/2 < 1/2$, and thus the event $x = 1$ is relatively rarer than $x = 0$. The next result states that at the optimum, the agent is relatively more likely to discard the more common event $x = 0$ in agreement with Kahneman and Tversky (1979), who observe that agents tend to overweight rare events.

Corollary 4. *At the optimum, the agent is biased in favor of the rare event* $x = 1$*:* $\beta_1^* > \beta_0^* > 0$ *(and her guess of the state equals the observed signal realization, i.e.* $\sigma = \sigma_I$*).*

To illustrate the result, consider an agent who is forming her probability belief over a flight accident. The accident probability per flight in the safe state of the world is 10^{-6}, whereas it is 10^{-5} in the dangerous state of the world, and both states are a priori equally likely. The agent can sequentially observe arbitrarily many past flight outcomes, but cannot aggregate the information,

[7] We can always achieve this by relabeling the states θ and the signals values x, unless $\rho_0 = \rho_1$ or $\rho_0 = 1 - \rho_1$.

and recalls only the last observed flight. She guesses that the state of the world is dangerous if and only if the last observed flight is eventful.

Consider first an agent who always terminates right after the observation of the first data-point. Such an agent benefits from a "second thought" whenever she observes an uneventful flight: Either the second observed flight will be uneventful, in which case the second thought will have been inconsequential, or the redrawn flight will be eventful and the agent will switch her assessment from the safe to the dangerous state. Such a switch is beneficial since conditional on two contradicting data-points the dangerous state is relatively more likely. More generally, conditional on two contradicting data-points, the state θ for which ρ_θ is closer to $1/2$ is relatively more likely, since it is relatively likely that such a state generates contradicting signal values. Thus, relative to the immediate termination strategy, the agent will benefit from discarding the uneventful flight observations with positive probability.

For the assumed accident rates, the optimal ratio of the search intensities is $\beta_1^*/\beta_0^* \approx 316,000$ where the signal value $x = 1$ stands for the eventful flight observation. Thus, the agent's search for data is heavily biased towards the rare accidents. This strategy generates probabilities of the correct state identification approximately equal to 0.76 in both states of the world. Since the immediate termination strategy identifies the correct state with a probability equal approximately to half, the bias significantly improves the payoff.

5.4 Salience

Bordalo, Gennaioli, and Shleifer (2012) interpret salience as directed attention focus. They quote the popular work by Daniel Kahneman (2011):

> "Our mind has a useful capability to focus on whatever is odd, different or unusual."

The quote states a causal relation between the two features of the salient phenomena: These are (i) odd, different or unusual, and because of (i), people benefit from (ii) focusing their attention on such phenomena. Here, we confirm Kahneman's intuition within our proposed framework. Our microfoundation of the salience effect is related to the insight emerging in psychological research on visual salience. Itti (2007) conceptualizes the visual salience effect as attention allocation to a subset of the visual field that is "sufficiently different from its surroundings to be worthy of [one's] attention." Similarly, in our model, a payoff state is salient if it stands out sufficiently from similar states to be worthy of the focus of the agent's information search.

As before, the agent faces a perceptual task that requires her to announce a realization of the random state θ. She is endowed with a primitive perception technology that generates a perceived value θ' of the state. The primitive perception is informative but noisy: the perceived value θ' equals the true state θ with a high probability, but mistakes, $\theta' \neq \theta$, occur sometimes. We view the primitive perception technology as a black-box model of a physiological sensor that generates a noisy first impression θ' of the true state θ. The agent can use the sensor repeatedly but is not able to aggregate the information. She conditions the repetition of the sensor's use on the most recent perception and announces the terminal perception.

We formalize this perception task as follows. The agent makes an announcement $a \in A = \Theta$, where $2 < |\Theta| < \infty$, and receives payoff $u(a, \theta) = 1$ if her announcement is correct, $a = \theta$, and $u(a, \theta) = 0$ if $a \neq \theta$. The prior is uniform. Each use of the agent's sensor generates a signal value/perception $\theta' \in X = \Theta$, with conditional probabilities $p(\theta' \mid \theta)$. We make two assumptions on p:

Symmetry: $p(\theta' \mid \theta) = p(\theta \mid \theta')$.

Sufficient precision: $p(\theta \mid \theta) > p(\theta' \mid \theta)$ for all $\theta \neq \theta'$.

For two states θ_1 and θ_2, we say that θ_1 is *more distinct* than θ_2 if for each state $\theta_3 \neq \theta_1, \theta_2$, $p(\theta_1 \mid \theta_3) < p(\theta_2 \mid \theta_3)$. Suppose for illustration that the perceptual task involves recognition of a color from a set {azure, indigo, red}. Intuitively, the red color stands out of this set, and this is captured by the above definition. Assume that the two shades of blue are similar in that the agent's first impression confuses them in 10% of cases, $p(\text{azure} \mid \text{indigo}) = p(\text{indigo} \mid \text{azure}) = 0.1$, but $p(\theta \mid \text{red}) = p(\text{red} \mid \theta) = 0.01$ for $\theta \in \{\text{azure}, \text{indigo}\}$. Then, the red color is more distinct according to our definition than either of the two blue shades.

We make a simplifying assumption that the agent uses the identity action strategy σ_I; she announces the state equal to her last perception. We also make a regularity assumption that the optimal termination probabilities β_x are positive for all $x \in \Theta$.[8] Let $r^* = r(p, \beta^*, \sigma_I)$ be the optimal feasible choice rule.

Proposition 3. *If state θ_1 is more distinct than state θ_2, then the agent's terminal perception is biased in favor of the more distinct state θ_1 at the expense of the less distinct state θ_2:*

$$r^*(\theta_1 \mid \theta_2) > r^*(\theta_2 \mid \theta_1).$$

Since the primitive perception technology p is symmetric by assumption, the asymmetry in favor of the distinct state of the optimal terminal perception r^* is driven solely by the optimization of the termination strategy. To gain the intuition for the salience of the distinct states, consider a state θ^* that is similar to many other states and an agent who always terminates the process after the first round: $\beta = \mathbf{1}$. This agent is relatively uninformed whenever she forms perception θ^*, since the true state differs from θ^* with a sizeable probability. The agent with this indistinct perception θ^* would thus benefit from "having a second thought"— i.e., from running the primitive perception formation process once again. The optimal termination strategy involves repeating the primitive process with relatively high probability whenever the agent forms a perception of an indistinct state, and this shifts the terminal perception in favor of the distinct states.

[8]This regularity assumption must be satisfied when $p(\theta \mid \theta)$ is sufficiently close to one for each θ, since then any strategy for which some β_x is zero is dominated by the strategy that immediately terminates and announces the first impression.

6 Sophisticated and impatient agents

We explore here the robustness of our results to an increase in the memory capacity of the agent and to discounting. The first subsection returns to Examples 2 and 3 from Section 2.2 in which the agent is endowed with more elaborate memory and can aggregate several observed signal realizations. We show that the two examples are special cases of our baseline model once we modify the definition of the signals and appropriately define the set $\mathcal{P}$ of the feasible experiments. Hence, the optimal rules solving the examples are second-thought-free, which in turn lead to the desirability of an asymmetric treatment of signals. The second subsection accommodates discounting.

6.1 Sophisticated decision processes

Examples 2 and 3 from Section 2.2 seemingly violate our baseline model in that the agent can aggregate information across signal realizations. We show, however, that these examples are formally special cases of our baseline model appropriately specified.

The formal specification of Example 2 (imperfect information aggregation) follows. The agent is endowed with one Blackwell experiment $\mu(x \mid \theta)$ with a finite signal space X and, additionally, with a finite set M of the memory states m. After each run of the experiment μ, the agent either terminates or continues with decision-making. If the agent continues, then she transitions from the current memory state to a new memory state and reruns the statistical experiment $\mu(x \mid \theta)$. That is, the agent selects a (generalization of the) termination strategy: $\gamma : M \times X \longrightarrow \Delta(M \cup \{\mathfrak{t}\})$, where $\gamma(m' \mid m, x)$ is the probability that the agent in the memory state m who has observed signal realization x in the last run of the experiment μ continues with the decision-making and transitions to the memory state m', and $\gamma(\mathfrak{t} \mid m, x)$ is the probability that such an agent terminates. The terminating agent chooses action $\sigma(m, x)$ that depends both on the current memory state and on the signal realization observed in the last run of μ. The agent starts the decision-making in the memory state m_0. A pair γ, σ induces a θ-dependent Markov chain over the memory states that eventually terminates with choice $\sigma(m, x)$, where m is the last memory state and x is the last signal realization observed. Let $p(a \mid \theta; \gamma, \sigma)$ be the probability that the agent terminates with the choice a in state θ, and let $\mathcal{P}_{iia}$ be the set of all stochastic choice rules p that this agent can construct. She selects the choice rule from $\mathcal{P}_{iia}$ that maximizes her ex ante expected payoff.

We now demonstrate that Example 2 is a special case of our baseline model. Consider the baseline model with the signal space $X = A$ and the set of the feasible primitive experiments $\mathcal{P} = \mathcal{P}_{iia}$. The set $\mathcal{R}(\mathcal{P}_{iia}) = \{r(p, \beta, \sigma) : p \in \mathcal{P}_{iia}, \beta \in B, \sigma \in S\}$ is then a set of the stochastic choice rules that can be constructed as follows. The agent runs any process $p \in \mathcal{P}_{iia}$, and observes a signal value/action recommendation a with probability $p(a \mid \theta)$. She terminates with probability β_a, according to the termination strategy $\beta = (\beta_a)_{a \in A}$, and upon the termination chooses an action $a' = \sigma(a)$, where $\sigma \in S$ is any mapping $A \longrightarrow A$. She reruns the process p with probability $1 - \beta_a$, observes a new action recommendation generated by p, et cetera, until she terminates after a stochastic number of repetitions of the process p.

As it turns out, no new choice rules beyond those from $\mathcal{P}_{iia}$ can be constructed by these selective repetitions. This follows because the repetitions of the rule $p \in \mathcal{P}_{iia}$ with the termination strategy β can always be replicated with an appropriate choice of a different rule in $\mathcal{P}_{iia}$ that whenever p would terminate with a restarts the process from scratch with probability $1 - \beta_a$. Formally:

Lemma 3. $\mathcal{R}(\mathcal{P}_{iia}) = \mathcal{P}_{iia}$.

According to the lemma, Example 2 is the special case of our baseline model with $\mathcal{P} = \mathcal{P}_{iia}$ and $X = A$, since in such a specification of the baseline model, the set of the feasible rules coincides with those in Example 2. In particular, the optimal choice rule $p^* \in \mathcal{P}_{iia}$ solving Example 2 coincides with the optimal rule $r^* \in \mathcal{R}(\mathcal{P}_{iia})$ solving this specification of the baseline model.

The repeated-cognition problem with $\mathcal{P} = \mathcal{P}_{iia}$ is purely formal in that the optimal termination probabilities $\beta^*_x = 1$ for all $x \in X = A$, and thus the agent conducts the optimal process $p^* \in \mathcal{P}_{iia}$ only once and terminates. Nevertheless, the observation that p^* solves the repeated-cognition problem has an important implication.

Corollary 5. *The choice rule that solves Example 2 (imperfect information aggregation) is second-thought-free.*

Wilson (2014) differs from this example mainly in that she assumes exogenous termination probabilities. By adding optimization over the terminations to the model of Wilson, we gained the partial characterization of the optimal choice rule without fully solving the problem: One can conclude that the optimal choice rule is second-thought-free without analyzing the optimal use of the memory states.

The same argument applies to Example 3 (partial forgetting). We first formalize Example 3 as follows. Let H be the set of the signal histories h of length $|h| \leq N$. The agent at a history h can (i) terminate her decision-making; (ii) discard some of the information accumulated; or (iii), if $|h| < N$, acquire a new signal realization. (i) An agent terminating at h chooses action $\sigma(h)$. (ii) An agent who discards some information transitions to a truncation h' of her current history h.[9] (iii) An agent who acquires a new signal realization transitions to a history hx, where x is the new signal realization drawn from $\mu(x \mid \theta)$. The decision-making is governed by a pair of mappings $\gamma : H \times \Theta \longrightarrow \Delta(H \cup \{\mathfrak{t}\})$ and $\sigma : H \longrightarrow A$, where $\gamma(h' \mid h, \theta)$ stands for the probability that the agent at history h in state θ continues decision-making and transitions to h', and $\gamma(\mathfrak{t} \mid h, \theta)$ is the probability of termination at history h in state θ. The mapping γ is constrained to satisfy 1. $\gamma(h' \mid h, \theta)$ is independent of θ if h' is a truncation of h, 2. $\gamma(\mathfrak{t} \mid h, \theta)$ is independent of θ, 3. $\frac{\gamma(hx|h,\theta)}{\gamma(hx'|h,\theta)} = \frac{\mu(x|\theta)}{\mu(x'|\theta)}$, 4. $\gamma(h' \mid h, \theta) = 0$ unless h' is a truncation of h, or $h' = hx$ for some $x \in X$ and $|hx| \leq N$. Constraints 1 and 2 require the agent to condition the decision to discard information or to terminate only on her current history independently of the state. Constraint 3 allows the agent to expand her information set only by running the experiment $\mu(x \mid \theta)$. Constraint 4 restricts each step of information acquisition to one draw from $\mu(x \mid \theta)$ or to a partial discarding

[9] A truncation is obtained by deleting one or more last elements in h.

of the accumulated information. Let $p(a \mid \theta; \gamma, \sigma)$ be the probability that the agent who employs (γ, σ) terminates with action a in the state θ. The agent chooses γ and σ to maximize her ex ante expected payoff.

As with the previous example, let $\mathcal{R}(\mathcal{P}_{pf})$ be the set of the feasible choice rules in our baseline model with the set of the feasible primitive experiments $\mathcal{P}$ identified with $\mathcal{P}_{pf}$.

Lemma 4. $\mathcal{R}(\mathcal{P}_{pf}) = \mathcal{P}_{pf}$.

Thus, again, the rule $p^* \in \mathcal{P}_{pf}$ solving Example 3, and the optimal rule $r^* \in \mathcal{R}(\mathcal{P}_{pf})$ coincide, and thus the rule solving the example must be second-thought-free.

Corollary 6. *The choice rule that solves Example 3 (partial forgetting) is second-thought-free.*

Given that the second-thought-free conditions are satisfied with sophisticated agents from Examples 2 and 3, and that some form of selective information processing is needed to meet those conditions, we conclude that the optimality of biases in the decision process will apply quite generally. Moreover, when the action and state sets are binary (while the signal space can be arbitrary) as considered in Section 4, then Proposition 2 derived above for unit-memory agents applies to the sophisticated agents from Examples 2 and 3 with a parameter $\overline{d}$ set to the maximal perceptual distance among the rules p in $\mathcal{P}_{iia}$ and $\mathcal{P}_{pf}$, respectively. In fact, a researcher who observes only the choice data and is ignorant of the decision process employed by the agent, cannot distinguish the simple agent with unit memory from her sophisticated counterparts from Examples 2 and 3. For such a researcher, the perceptual distance $\overline{d}$ is a free single-valued parameter of the model, and Proposition 2 predicts stochastic choice behavior as a function of $\overline{d}$, of the prior and the incentives.

6.2 Costly delay

Our baseline model abstracts from the cost of time in that the agent is only concerned with how the repetitions of the signal extraction affect the correlation of the signal with the state. We now incorporate discounting.

We continue to study the baseline model from Section 2.1, except that the agent discounts future payoffs exponentially with the discount factor $\delta \in (0, 1)$. To accommodate discounting, we redefine the choice rule induced by the experiment p, the termination strategy β and the action strategy σ as follows.

$$r_\delta(a \mid \theta; p, \beta, \sigma) = \sum_t \sum_{\mathbf{x}^t : \sigma(x_t) = a} \delta^t \rho\left(\mathbf{x}^t \mid \theta; p, \beta\right), \tag{12}$$

where $\rho\left(\mathbf{x}^t \mid \theta; p, \beta\right)$ defined in (1) is the conditional probability of the signal history $\mathbf{x}^t$. That is, $r_\delta(a \mid \theta; p, \beta, \sigma)$ is the *discounted* probability of the choice of action a in the state θ. When $\delta = 1$, then (12) coincides with our baseline definition of the choice rule.

The set of the feasible discounted rules is $\mathcal{R}_\delta(\mathcal{P}) = \{r_\delta(p, \beta, \sigma) : p \in \mathcal{P}, \beta \in B, \sigma \in S\}$. The discounted repeated-cognition problem is to select a feasible rule r_δ that maximizes the expected

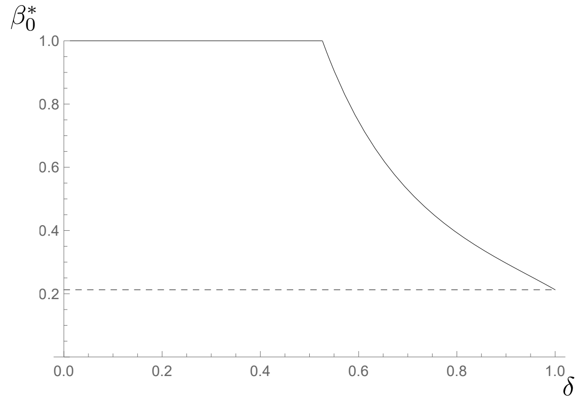

Figure 2: Confirmation bias with discounting. Action 1 is a priori more attractive: $\pi_1 u_1 = 10\times\pi_0 u_0$. The primitive experiment is symmetric: $p(1 \mid 1) = p(0 \mid 0) = .9$. The agent terminates immediately when she observes signal value 1, $\beta_1^* = 1$. When $\delta > .53$, then the agent is biased towards state 1: when she encounters signal value 0, then she terminates the decision-process with a probability only $\beta_0^*(\delta) < 1$ (the full curve). The dotted line is β_0^*/β_1^* from the baseline model without discounting.

payoff:

$$\max_{r_\delta \in \mathcal{R}_\delta(\mathcal{P})} \sum_{\theta\in\Theta, a\in A} \pi_\theta r_\delta(a \mid \theta) u(a,\theta), \tag{13}$$

where discounting is incorporated in the definition of the feasible rules.

The next result generalizes the second-thought-free condition. Let $r_\delta^* = r_\delta(p^*, \beta^*, \sigma^*)$ be the choice rule solving the discounted repeated-cognition problem (13).

Proposition 4. *If the termination strategy $\beta_x^* \in (0,1)$ is interior for all x such that $\sigma^*(x) = a$, then*

$$\sum_{\theta\in\Theta} \pi_\theta u(a,\theta) r_\delta^*(a \mid \theta) = \delta \sum_{\theta\in\Theta, a'\in A} \pi_\theta u(a',\theta) r_\delta^*(a' \mid \theta) r_\delta^*(a \mid \theta). \tag{14}$$

The condition has the same interpretation as the second-thought-free condition in the absence of discounting. The left-hand side is the payoff for following the optimal decision process r_δ^* summed up across all contingencies that terminate with choice of a. The right-hand side is the payoff that the agent would get across the same contingencies if she restarted the decision process r_δ^* instead of the termination.

For illustration, we now revisit the confirmation bias application from Section 5.1 with an impatient agent. We find that, unless discounting is too strong, the impatient agent chooses qualitatively the same strategy as the patient one, although the impatient agent speeds up her decision-making by choosing larger termination probabilities.

The setting is as follows. The agent chooses $a \in \{0,1\}$ and receives $u(a,\theta) = u_\theta > 0$ if $a = \theta$, and zero reward otherwise. Action 1 is a priori more attractive than action 0; $\pi_1 u_1 > \pi_0 u_0$. The agent has access to a single primitive experiment p that generates signal values in $X = \{0,1\}$. The experiment is symmetric with probabilities $p(1 \mid 1) = p(0 \mid 0) = \alpha > 1/2$.

Proposition 5. *The agent chooses the action equal to the last observed signal realization. She terminates her decision-making immediately after she encounters signal realization* 1*:* $\beta_1^* = 1$*. When* $\delta \in \left(\frac{1}{\alpha+(1-\alpha)R}, 1\right]$*, then the agent who observes* $x = 0$ *terminates with interior probability* $\beta_0^* \in (0,1)$ *that decreases in* δ*. When* $\delta \in \left(0, \frac{1}{\alpha+(1-\alpha)R}\right)$*, then the agent terminates immediately:* $\beta_0^* = \beta_1^* = 1$*.*

7 Discussion of empirical evidence

We briefly discuss the empirical evidence in relation to the main predictions of our model in light of the state recognition problem that has been widely studied by psychologists. Our main predictions are in line with Ratcliff and McKoon (2008). These psychologists asked lab participants to report a binary state visually encoded by a pattern of moving dots on a computer screen. We conceptualize the sequence of the observed dot movements as a sequence of the signal values generated by a primitive experiment p in our model. Ratcliff and McKoon's design exhibits a symmetry that justifies the symmetry $p(1 \mid 1) = p(0 \mid 0)$ assumed in our primary example from Sections 5.1 and 5.2. The experiment included two treatments with non-uniform prior. The prior in each of these treatments was announced to participants in instructions, and it was symmetric with the more likely state permuted across the two treatments.

The data exhibit the following two styled facts:

1. The posterior probability that the participant's announcement is correct is higher when she announces the a priori expected state than when she makes the surprising announcement.

2. Late announcements are relatively less precise.

Both these stylized facts are in line with our predictions. The next corollary of Proposition 2 states that posterior beliefs of our agent increase monotonically with the prior belief.

Corollary 7. *Let* $R \in (1/\overline{d}, \overline{d})$*, and fix rewards* u_θ*. Posterior probabilities of the state 1,* $\Pr_{r^*}(\theta = 1 \mid a = 1)$ *and* $\Pr_{r^*}(\theta = 1 \mid a = 0)$*, increase with the prior probability* π_1 *of the state 1.*

Consider, as in the experiment, the case with symmetric incentives, $u_1 = u_0$. Since, by symmetry, $\Pr_{r^*}(\theta = 1 \mid a = 1)$ for prior π_1 equals $\Pr_{r^*}(\theta = 0 \mid a = 0)$ for the prior $\pi_1' = 1 - \pi_1$, it follows that when the agent announces the a priori likely state then her announcement is more likely correct than the surprising announcement, which matches the stylized fact 1. The stylized fact 2 is in line with our subsection 5.2.

The two stylized facts are at odds with the predictions of the influential sequential-learning model of Wald (1945). Wald's agent faces an opportunity cost of time, and learns about a binary state from a sequence of weakly informative signals. The optimal decision procedure terminates once the agent reaches one of the two absorbing posterior beliefs. The optimal absorbing posteriors are independent of the agent's prior (as long as the prior is not too extreme). When the incentives are symmetric, as in Ratcliff and McKoon's experiment in which participants were simply asked to announce the correct state, the two absorbing posteriors have the same precision, which fails to match both stylized facts.

Fudenberg et al. (2018) have recently generalized Wald's model by adding an uncertainty about the size of the stake to the sequential-sampling problem. The speed-accuracy complementarity arises in their model since their agent interprets long spells of inconclusive learning as an indication that the stake is small, thereby leading her to optimally terminate the process with imprecise posteriors. The authors' proposed mechanism differs from ours. While their agent "gives up" after long delays, in our model long delays indicate that the agent has encountered the surprising state for which her decision process is poorly adapted. Since there is no uncertainty about the magnitude of the stakes in Ratcliff and McKoon's experiment, the model of Fudenberg et al. is not readily applicable to the discussed data.[10]

8 Summary

Agents, who cannot comprehend all facts available to them, benefit from selective attention. We show that agents can implement a targeted information search in a process that resembles the natural phenomenon of hesitation. Like a hesitant person, the agent can, conditional on the action contemplated, decide whether she implements the action or whether she will have a second thought, and run the cognition process once more. Such hesitation can be productive, despite consisting of repetitions of the same stochastic cognition process. By conditioning the probability of the repetition on the conclusion of the reasoning, the agent controls the correlation of her terminal conclusion and the payoff state. The optimal decision process arising in our model exhibits natural hesitation patterns: The agent will have second thoughts—that is, she will repeat her cognition—whenever the expected payoff for the currently favored choice is inferior to the expected payoff for continuing decision-making. At optimum, the agent terminating the decision-making must be indifferent between terminating with the currently contemplated action, and repeating the process. In a sense, the condition formalizes the concept of a reasonable doubt. Abstracting from many considerations such as information aggregation across the jury members, a jury deciding a trial under common law should be, if using the optimal decision procedure, indifferent between declaring a verdict and announcing a hung jury and initiate retrial.

Let us conclude by reviewing the limitations of our main result. The central assumption—the ability of the agent to freely repeat her decision process—may fail for several reasons. One reason is that the agent may only have access to a limited data set that constrains her to a finite number of repetitions of the primitive decision process, making the optimal termination strategy non-stationary. Another complication arises if the outcomes of distinct runs of the same cognition process are not conditionally independent as assumed in our model; this may arise if some

[10] Che and Mierendorff (2016) consider learning about a binary state from evidence arriving at a Poisson rate. They obtain for some parameter specifications that conditional on one particular action being made, the probability that this action is correct declines with time. Their result does not aggregate the mistake rates over all possible actions, making the connection to Ratcliff and McKoon findings less clear.

cognition errors are systematic and are likely to emerge in distinct repetitions of the cognition. We conjecture that the second-thought-free condition holds in such a case, with the agent internalizing the correlations between the cognition runs.

A Proofs

A.1 Proofs for Section 3

Proof of Lemma 1. Suppose by contradiction that (4) holds with strict inequality for some a chosen with positive probability. Then, $\mathrm{E}_\alpha\left[\mathrm{E}_\alpha\left[u(a_1,\theta)\mid a_1\right]\right] > \mathrm{E}_\alpha\left[\mathrm{E}_\alpha\left[u(a_2,\theta)\mid a_1\right]\right]$, and applying the law of iterated expectation, this simplifies to $\mathrm{E}_\alpha\left[u(a_1,\theta)\right] > \mathrm{E}_\alpha\left[u(a_2,\theta)\right]$. This establishes the contradiction since a_1 and a_2 are conditionally iid. draws generated by the rule r. □

Proof of Lemma 2. The effective experiment $s(p,\beta)$ satisfies a recursion

$$s(x\mid\theta;p,\beta)=\beta_x p(x\mid\theta)+\sum_{x'\in X}\left(1-\beta_{x'}\right)p\left(x'\mid\theta\right)s(x\mid\theta;p,\beta),$$

where the first summand is the probability that the agent terminates with signal value x after the first run of the primitive experiment p, and the second summand is the probability that the agent continues with decision-making after the first run of p and terminates with x later. Solving for $s(x\mid\theta;p,\beta)$ gives (6). □

Proof of Proposition 1. Let (p^*,β^*,σ^*) solve the repeated-cognition problem. Consider an action a chosen with a positive probability. There must exist x such that $\sigma^*(x)=a$ and $\beta^*_x>0$. Therefore, the constraint $\beta_x\geq 0$ is not binding for this x, and the first-order condition of problem (7) with respect to β_x is:

$$\begin{aligned}\sum_{\theta\in\Theta}\pi_\theta\frac{s\left(x\mid\theta;p^*,\beta^*\right)}{\beta^*_x}u(a,\theta)-\sum_{\theta\in\Theta,x'\in X}\pi_\theta s\left(x'\mid\theta;p^*,\beta^*\right)\frac{s\left(x\mid\theta;p^*,\beta^*\right)}{\beta^*_x}u\left(\sigma^*(x'),\theta\right) &= \\ \sum_{\theta\in\Theta}\pi_\theta\frac{s\left(x\mid\theta;p^*,\beta^*\right)}{\beta^*_x}u(a,\theta)-\sum_{\theta\in\Theta,a'\in A}\pi_\theta r\left(a'\mid\theta;p^*,\beta^*,\sigma^*\right)\frac{s\left(x\mid\theta;p^*,\beta^*\right)}{\beta^*_x}u\left(a',\theta\right) &\geq 0,\end{aligned}$$

where we have summed over all x' such that $\sigma^*(x')=a'$ in the second line. Multiplication by β^*_x and summation over all x such that $\sigma^*(x)=a$ and $\beta_x>0$ give

$$\sum_{\theta\in\Theta}\pi_\theta r\left(a\mid\theta;p^*,\beta^*,\sigma^*\right)u(a,\theta)-\sum_{\theta\in\Theta,a'\in A}\pi_\theta r\left(a'\mid\theta;p^*,\beta^*,\sigma^*\right)r\left(a\mid\theta;p^*,\beta^*,\sigma^*\right)u(a',\theta)\geq 0.$$

Division of this by $\sum_\theta \pi_\theta r\left(a\mid\theta;p^*,\beta^*,\sigma^*\right)$ leads to (4). So far, we have proved that the terminating agent weakly prefers termination to continuation. Lemma 1 implies (5)—the indifference between termination and continuation. □

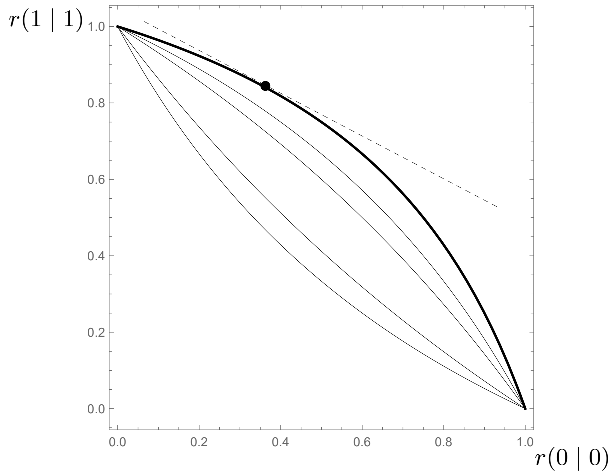

Figure 3: Each point in $[0,1]^2$ on this graph corresponds to a choice rule. Each depicted curve is a set $\mathcal{R}_{p,x_0,x_1}$ of the choice rules constructible from an experiment p and a pair of the signal values x_0, x_1. The thick curve corresponds to the experiment $\overline{p}$ and the signal pair that maximizes the perceptual distance in (9). Since the objective—the expected utility maximization—is linear in the choice rule, the indifference curves are downward sloping lines. The dashed line is the tangential indifference curve. The dot depicts the solution of the repeated-cognition problem.

A.2 Proofs for Section 4

Proof of Proposition 2. Let us fix a statistical experiment p and an ordered pair x_0, x_1 of signal realizations from X. Our first step characterizes the feasibility set $\mathcal{R}_{p,x_0,x_1}$ of all choice rules $r(a \mid \theta)$ that the agent can construct by (i) termination strategies that have zero termination probability for all signal values other than x_0, x_1, and by (ii) the action strategy $\sigma(x_0) = 0$ and $\sigma(x_1) = 1$. Let $d_{p,x_0,x_1} = \frac{p(x_1|1)p(x_0|0)}{p(x_0|1)p(x_1|0)}$. We claim that

$$\mathcal{R}_{p,x_0,x_1} = \{r : r(1 \mid 1)r(0 \mid 0) = d_{p,x_0,x_1} r(1 \mid 0)r(0 \mid 1)\}. \tag{15}$$

That is, a rule r can be constructed if and only if it preserves $\frac{r(1|1)r(0|0)}{r(0|1)r(1|0)} = d_{p,x_0,x_1}$, or if it always selects a same action. By controlling the termination probabilities β_{x_0}, and β_{x_1} the agent trades off the likelihoods $r(0 \mid 0; p, \beta, \sigma)$ and $r(1 \mid 1; p, \beta, \sigma)$ of the correct choice in the states 0 and 1, respectively. See Figure 3. The set $\mathcal{R}_{p,x_0,x_1}$ of the rules constructible from any given p and a pair of signal values is compact.

We now show (15). For any β such that $\beta_{x'} = 0$ for $x' \neq x_0, x_1$ and $\beta_{x'} > 0$ for $x' = x_0, x_1$ we have

$$\frac{r(1 \mid 1; p, \beta, \sigma)r(0 \mid 0; p, \beta, \sigma)}{r(0 \mid 1; p, \beta, \sigma)r(1 \mid 0; p, \beta, \sigma)} = \frac{\frac{\beta_{x_1} p(x_1|1)}{\sum_{x \in \{x_0, x_1\}} \beta_x p(x|1)} \frac{\beta_{x_0} p(x_0|0)}{\sum_{x \in \{x_0, x_1\}} \beta_x p(x|0)}}{\frac{\beta_{x_0} p(x_0|1)}{\sum_{x \in \{x_0, x_1\}} \beta_x p(x|1)} \frac{\beta_{x_1} p(x_1|0)}{\sum_{x \in \{x_0, x_1\}} \beta_x p(x|0)}} = \frac{p(x_1 \mid 1)p(x_0 \mid 0)}{p(x_0 \mid 1)p(x_1 \mid 0)} = d_{p,x_0,x_1}.$$

Thus, every $r \in \mathcal{R}_{p,x_0,x_1}$ either always selects a same action, or it selects both actions with a positive probability and satisfies $\frac{r(1|1)r(0|0)}{r(0|1)r(1|0)} = d_{p,x_0,x_1}$. Vice versa, if a rule r' satisfies $\frac{r'(1|1)r'(0|0)}{r'(0|1)r'(1|0)} = d_{p,x_0,x_1}$, then it belongs to $\mathcal{R}_{p,x_0,x_1}$. To see this, first observe that there exists $\alpha = \beta_{x_1}/\beta_{x_0}$ such that

$$r'(1 \mid 1) = \frac{\beta_{x_1} p(x_1 \mid 1)}{\beta_{x_0} p(x_0 \mid 1) + \beta_{x_1} p(x_1 \mid 1)} = \frac{\alpha p(x_1 \mid 1)}{p(x_0 \mid 1) + \alpha p(x_1 \mid 1)}. \quad (16)$$

Second, we need to prove that the same value of α satisfies

$$r'(0 \mid 0) = \frac{p(x_0 \mid 0)}{p(x_0 \mid 0) + \alpha p(x_1 \mid 0)}.$$

This is indeed the case since $r'(0 \mid 0)$ solves $\frac{r'(1|1)r'(0|0)}{(1-r'(1|1))(1-r'(0|0))} = d_{p,x_0,x_1}$, and hence

$$r'(0 \mid 0) = \frac{d_{p,x_0,x_1}\left(1 - r'(1 \mid 1)\right)}{d_{p,x_0,x_1}\left(1 - r'(1 \mid 1)\right) + r'(1 \mid 1)} = \frac{p(x_0 \mid 0)}{p(x_0 \mid 0) + \alpha p(x_1 \mid 0)},$$

as needed. The second equality follows from substituting the right-hand side of (16) for $r'(1 \mid 1)$ and the definition of d_{p,x_0,x_1}.[11] This concludes the proof of (15).

In the second step, we consider an agent who can choose an experiment $p \in \mathcal{P}$, a pair of signals $x_0, x_1 \in X$, is constrained to set $\beta_{x'}$ to 0 for all $x' \neq x_0, x_1$ and must use the action strategy $\sigma(x_a) = a$. We will show that this constraint is not binding in the third step of the proof. Note that the problem of this constrained agent has a solution since this agent chooses a rule from a finite collection of the compact sets $\mathcal{R}_{p,x_0,x_1}$ and her objective is continuous in r.

Such constrained agent either chooses a rule r that always selects a same action or chooses p, x_0, x_1 that maximizes d_{p,x_0,x_1}. To see this, fix any value $r(0 \mid 0) \in (0,1)$, let $r(1 \mid 1)$ be given by $\frac{r(0|0)r(1|1)}{(1-r(0|0))(1-r(1|1))} = d$, and observe that $r(1 \mid 1)$ increases in d. Therefore, the optimal rule lies in the set $\mathcal{R}^* = \left\{r : r(1 \mid 1)r(0 \mid 0) = \overline{d}r(1 \mid 0)r(0 \mid 1)\right\}$. Since the agent's objective is linear in r, the optimal rule is the point of tangency of $\mathcal{R}^*$ and of an indifference line; see Figure 3. The slope $\frac{\partial r(0|0)}{\partial r(1|1)}$ for $r \in \mathcal{R}^*$ is decreasing in $r\,(1 \mid 1)$ and attains value $-1/\overline{d}$ for $r\,(1 \mid 1) = 0$, and value $-\overline{d}$ for $r\,(1 \mid 1) = 1$. Thus, when $R < 1/\overline{d}$ or $R > \overline{d}$, then the problem has the corner solution as specified in statements 1 and 2 of the proposition.

When $R \in \left(1/\overline{d}, \overline{d}\right)$, then the optimal choice rule r^* satisfies the feasibility condition $\frac{r^*(1|1)r^*(0|0)}{r^*(0|1)r^*(1|0)} = \overline{d}$, the second-thought-free condition (5) (applied to action $a = 1$):

$$\pi_1 u_1 r^*(1 \mid 1) = \pi_0 u_0 r^*(0 \mid 0) r^*(1 \mid 0) + \pi_1 u_1 r^*(1 \mid 1) r^*(1 \mid 1),$$

and two normalization conditions $\sum_a r^*(a \mid \theta) = 1$, for $\theta = 0, 1$. These four conditions jointly imply the explicit solution for the optimal choice rule in (10). The expression (11) for β_1^*/β_0^* follows from (10) and the condition $\frac{r^*(1|\theta)}{r^*(0|\theta)} = \frac{\beta_1^* \overline{p}(1|\theta)}{\beta_0^* \overline{p}(0|\theta)}$.

In the third step we prove that, if the solution of the repeated-cognition problem selects both

[11] A rule that always selects an action a can be trivially constructed by using $\beta_{x_a} = 1$ and $\beta_{x_{a'}} = 0$ for $a' \neq a$.

actions with positive probabilities, then it is optimal to use β_x that is positive for exactly two values of x.[12] Fix the primitive experiment p employed by the agent, let β be an optimal termination strategy for the given p, let X' be the set of signal values with positive β_x, and write concisely $s(x \mid \theta)$ for the effective experiment $s(x \mid \theta; p, \beta)$ induced by p and β. Let $s(x) = \sum_\theta \pi_\theta s(x \mid \theta)$ stand for the unconditional effective probability of x. For $x \in X'$ let $q_x \in \Delta(\Theta)$ be the posterior belief upon terminating with x: $q_x(\theta) = \pi_\theta s(x \mid \theta)/s(x)$.

We prove that the set of the posteriors $q_x(\theta)$ that the agent attains is binary. Suppose for contradiction that the agent attains three or more distinct posterior beliefs. Since the state space Θ is binary, there exists a signal value $x^* \in X'$ such that q_{x^*} is in the interior of the convex hull of the posteriors q_x, $x \in X' \setminus \{x^*\}$. Let μ_x be the coefficients that decompose q_{x^*} into q_x, $x \in X' \setminus \{x^*\}$. That is, $\mu \in \Delta\left(X' \setminus \{x^*\}\right)$ such that $q_{x^*}(\theta) = \sum_{x \in X' \setminus \{x^*\}} \mu_x q_x(\theta)$ for all $\theta \in \Theta$.

We will construct an alternative feasible effective experiment $\tilde{s}(x \mid \theta)$ with unconditional probabilities of x denoted by $\tilde{s}(x)$ and the posteriors $\pi_\theta \tilde{s}(x \mid \theta)/\tilde{s}(x)$ denoted by $\tilde{q}_x(\theta)$ such that:

$$\tilde{s}(x) = \begin{cases} s(x) + s(x^*)\mu_x \text{ if } x \in X' \setminus \{x^*\}, \\ 0 \text{ otherwise,} \end{cases} \tag{17}$$

and

$$\tilde{q}_x(\theta) = q_x(\theta) \text{ for all } x \in X' \setminus \{x^*\}, \theta \in \Theta. \tag{18}$$

Since the experiment $\tilde{s}$ is more informative than s (in the sense of the Blackwell comparison), the agent strictly prefers the alternative feasible effective experiment $\tilde{s}$ to s, as needed for the contradiction.

The construction of $\tilde{s}$ involves re-sampling whenever the agent encounters the signal value x^*. Note that if an effective experiment $s(x \mid \theta; p, \beta) = \frac{\beta_x p(x|\theta)}{\sum_{x' \in X} \beta_{x'} p(x'|\theta)}$ is induced by some p and β, then for any vector of probabilities $\tilde{\beta}_x$, the experiment

$$\tilde{s}(x \mid \theta) = \frac{\tilde{\beta}_x s(x \mid \theta; p, \beta)}{\sum_{x' \in X} \tilde{\beta}_{x'} s(x' \mid \theta; p, \beta)} = \frac{\tilde{\beta}_x \beta_x p(x \mid \theta)}{\sum_{x' \in X} \tilde{\beta}_{x'} \beta_{x'} p(x' \mid \theta)}$$

is also feasible, since it is induced by p and $\beta' = (\tilde{\beta}_x \beta_x)_{x \in X}$.

We claim that if

$$\tilde{\beta}_x = \begin{cases} c\left(1 + \frac{s(x^*)\mu_x}{s(x)}\right) \text{ if } x \in X' \setminus \{x^*\}, \\ 0 \text{ otherwise,} \end{cases}$$

where c is a constant such that $\tilde{\beta}_x \in (0, 1)$ for all $x \in X$, then the resulting $\tilde{s}$ satisfies the properties

[12] This insight exploits the assumption of perfect patience, since impatient agents would trade off informativeness against delay costs. We conjecture that when exponential discounting is considered, then the result that the agent ignores all but two signal realizations continues to hold for sufficiently patient agents and generic signal structures.

(17) and (18). Let us check:

$$
\begin{aligned}
\tilde{s}(x \mid \theta) &= \frac{\tilde{\beta}_x s(x \mid \theta)}{\sum_{x' \in X' \setminus \{x^*\}} \tilde{\beta}_{x'} s(x' \mid \theta)} \\
&= \frac{\tilde{\beta}_x s(x \mid \theta)}{c\left(\sum_{x' \in X' \setminus \{x^*\}} s(x' \mid \theta) + \sum_{x' \in X' \setminus \{x^*\}} \frac{s(x^*)\mu_{x'}}{s(x')} s(x' \mid \theta)\right)} \\
&= \frac{\tilde{\beta}_x s(x \mid \theta)}{c\left(\sum_{x' \in X' \setminus \{x^*\}} s(x' \mid \theta) + \sum_{x' \in X' \setminus \{x^*\}} \frac{s(x^*)\mu_{x'}}{\pi_\theta} q_{x'}(\theta)\right)} \\
&= \frac{\tilde{\beta}_x s(x \mid \theta)}{c\left(\sum_{x' \in X' \setminus \{x^*\}} s(x' \mid \theta) + \frac{s(x^*)}{\pi_\theta} q_{x^*}(\theta)\right)} \\
&= \frac{\tilde{\beta}_x s(x \mid \theta)}{c\left(\sum_{x' \in X' \setminus \{x^*\}} s(x' \mid \theta) + s(x^* \mid \theta)\right)} \\
&= \frac{\tilde{\beta}_x s(x \mid \theta)}{c} \\
&= \left(1 + \frac{s(x^*)\mu_x}{s(x)}\right) s(x \mid \theta).
\end{aligned}
$$

The property (17) holds since for all $x \in X' \setminus \{x^*\}$:

$$
\tilde{s}(x) = \left(1 + \frac{s(x^*)\mu_x}{s(x)}\right) s(x) = s(x) + s(x^*)\mu_x.
$$

To establish the property (18), check that for all $x \in X' \setminus \{x^*\}$ and all $\theta \in \Theta$:

$$
\tilde{q}_x(\theta) = \frac{\pi_\theta \tilde{s}(x \mid \theta)}{\sum_{\theta' \in \Theta} \tilde{s}(x \mid \theta')} = \frac{\pi_\theta \left(1 + \frac{s(x^*)\mu_x}{s(x)}\right) s(x \mid \theta)}{\sum_{\theta' \in \Theta} \left(1 + \frac{s(x^*)\mu_x}{s(x)}\right) s(x \mid \theta')} = \frac{\pi_\theta s(x \mid \theta)}{\sum_{\theta' \in \Theta} s(x \mid \theta')} = q_x(\theta).
$$

We have proven that if the agent chooses both actions with positive probabilities then she attains two posteriors. Additionally, since we have imposed the regularity condition that the maximizer in (9) is unique, β_x is positive for exactly two values of the signal x. Otherwise, there must exist distinct signals values that lead to a same posterior. We can then choose one signal value for each posterior and let the agent re-sample upon observing all other signal values. In this way, we can find distinct pairs of the signal values that implement the same set of optimal posteriors, which implies that d_{p,x_0,x_1} is maximized for distinct signal pairs, contradicting thus the regularity condition. □

Proof of Corollary 1. Statement 1 follows from the signs of the derivatives of the explicit solutions (10) and (11). For the relative response rate in statement 2, the monotone likelihood property of $\overline{p}$ implies that

$$
\frac{f_1}{f_0} = \frac{\beta_0^* \overline{p}(0 \mid 1) + \beta_1^* \overline{p}(1 \mid 1)}{\beta_0^* \overline{p}(0 \mid 0) + \beta_1^* \overline{p}(1 \mid 0)}
$$

increases in β_1^*/β_0^*, since

$$\frac{\partial \frac{f_1}{f_0}}{\partial \frac{\beta_1^*}{\beta_0^*}} = \frac{\overline{p}(1 \mid 1)\overline{p}(0 \mid 0) - \overline{p}(0 \mid 1)\overline{p}(1 \mid 0)}{\left(\overline{p}(0 \mid 0) + \frac{\beta_1^*}{\beta_0^*}\overline{p}(1 \mid 0)\right)^2} > 0.$$

□

A.3 Proofs for Section 5

Proof of Corollary 2. Since β_1^*/β_0^* increases in R, it suffices to show that $\beta_1^*/\beta_0^* = 1$ when $R = 1$ and the primitive experiment p is symmetric. Indeed, when $R = 1$, then by (11),

$$\frac{\beta_1^*}{\beta_0^*} = \sqrt{d_p}\frac{p(0 \mid 1)}{p(1 \mid 1)} = \sqrt{\frac{p(0 \mid 0)p(0 \mid 1)}{p(1 \mid 1)p(1 \mid 0)}} = 1,$$

where the last equality follows from the symmetry of p. □

Proof of Corollary 3. $\beta_1^* > \beta_0^*$ by Corollary (2) since $\pi_1 u_1 > \pi_0 u_0$. Recall that $f_\theta = \beta_1^* p(1 \mid \theta) + \beta_0^* p(0 \mid \theta)$ denotes the probability of termination per each round in state θ, and that the response time t in the state θ is geometrically distributed with the decision rate f_θ: $\Pr_\varphi(t \mid \theta) = f_\theta(1 - f_\theta)^t$ for $t = 0, 1, \ldots$. Since $p(1 \mid 1) = p(0 \mid 0) > p(1 \mid 0) = p(1 \mid 0)$ and $\beta_1^* > \beta_0^*$, the decision rate is higher in state 1 than in state 0: $f_1 > f_0$. Thus, the likelihood ratio $\Pr_\varphi(t \mid \theta = 1)/\Pr_\varphi(t \mid \theta = 0)$ decreases with t, and hence $\Pr_\varphi(\theta = 1 \mid t)$ decreases in t. The fact that $\beta_1^* > \beta_0^*$, and the symmetry of p implies that the probability of the correct choice is larger in state 1 than in state 0:

$$r\left(1 \mid 1; p, \beta^*, \sigma_I\right) = \frac{\beta_1^* p(1 \mid 1)}{\beta_0^* p(0 \mid 1) + \beta_1^* p(1 \mid 1)} > \frac{\beta_0^* p(0 \mid 0)}{\beta_0^* p(0 \mid 0) + \beta_1^* p(1 \mid 0)} = r\left(0 \mid 0; p, \beta^*, \sigma_I\right).$$

Since $\Pr_\varphi(a = \theta \mid t) = \Pr_\varphi(\theta = 1 \mid t) r(1 \mid 1; p, \beta^*, \sigma_I) + \Pr_\varphi(\theta = 0 \mid t) r(0 \mid 0; p, \beta^*, \sigma_I)$, the result obtains. □

Proof of Corollary 4. The belief formation problem studied is a special case of our binary setting with the primitive experiment $p(x \mid \theta) = \rho_\theta$ if $x = 1$, $p(x \mid \theta) = 1 - \rho_\theta$ if $x = 0$ and with equally a priori attractive actions, $R = 1$. Since $\rho_0 < \rho_1$, the labeling of the signals satisfies the monotone likelihood property. Since $R = 1 \in (1/d, d)$, Proposition 2 implies that the agent's behavior is stochastic, both β_0^* and β_1^* are positive, and the ratio of the search intensities β_1^*/β_0^* satisfies (11). Since $R = 1$, (11) simplifies to

$$\frac{\beta_1^*}{\beta_0^*} = d_p^{1/2}\frac{p(0 \mid 1)}{p(1 \mid 1)} = \left(\frac{p(0 \mid 1)p(0 \mid 0)}{p(1 \mid 1)p(1 \mid 0)}\right)^{1/2} = \left(\frac{(1 - \rho_1)(1 - \rho_0)}{\rho_1 \rho_0}\right)^{1/2}.$$

The inequality $\beta_1^* > \beta_0^*$ follows from $\rho_0 < \rho_1 < 1 - \rho_0$. □

The next result is an auxiliary lemma used in the proof of Proposition 3.

Lemma 5. *Suppose that termination probabilities β_x are positive for all $x \in \Theta$. Then, the optimal effective choice rule r^* satisfies for any pair of states θ, $\theta' \in \Theta$:*

$$r^*(\theta \mid \theta) r^*(\theta' \mid \theta) = r^*(\theta \mid \theta') r^*(\theta' \mid \theta'). \tag{19}$$

Condition (19) is a strengthening of the second-thought-free condition (5). It requires that the agent who has terminated the decision process with perception θ, and knows that the second run of the process r^* terminates with a value θ' is indifferent between θ and θ'. This condition is stronger than the second-thought-free condition (5), since (5) requires (19) to hold only on average across all θ'. This strengthening holds for the special case of this application with a symmetric experiment p.

Proof of Lemma 5. The optimal effective choice rule satisfies the second-thought-free condition (5), equivalent to:

$$r^*(\theta \mid \theta) = \sum_{\theta' \in \Theta} r^*(\theta \mid \theta') r^*(\theta' \mid \theta') \text{ for all } \theta \in \Theta,$$

which after two algebraic steps gives:

$$r^*(\theta \mid \theta)\big(1 - r^*(\theta \mid \theta)\big) = \sum_{\theta' \neq \theta} r^*(\theta \mid \theta') r^*(\theta' \mid \theta') \text{ for all } \theta \in \Theta,$$

$$\sum_{\theta' \neq \theta} r^*(\theta \mid \theta) r^*(\theta' \mid \theta) = \sum_{\theta' \neq \theta} r^*(\theta' \mid \theta') r^*(\theta \mid \theta') \text{ for all } \theta \in \Theta.$$

The last system of equations is formally equivalent to the system of balance conditions for a Markov chain. To see this, consider an ergodic Markov chain with transition probabilities from θ to θ' equal to $r^*(\theta' \mid \theta)$. The balance condition for the stationary distribution $\mu(\theta)$ of this chain is

$$\sum_{\theta' \neq \theta} \mu(\theta) r^*(\theta' \mid \theta) = \sum_{\theta' \neq \theta} \mu(\theta') r^*(\theta \mid \theta') \text{ for all } \theta \in \Theta,$$

and thus $r^*(\theta \mid \theta)$ is proportional to the ergodic probability $\mu(\theta)$ of the state θ for the chain with transition probabilities $r^*(\theta' \mid \theta)$.

Recall that if a Markov chain with transition probabilities $m(\theta' \mid \theta)$ is reversible, then its stationary distribution $\mu(\theta)$ satisfies detailed balance conditions

$$\mu(\theta) m(\theta' \mid \theta) = \mu(\theta') m(\theta \mid \theta') \text{ for all } \theta \neq \theta'.$$

Thus, it suffices to prove that the probabilities $r^*(\theta' \mid \theta)$ constitute a reversible Markov chain.

Recall that a Markov chain $m(\theta' \mid \theta)$ is reversible if and only if it satisfies the Kolmogorov criterion, which requires for all sequences of states $\theta_1, \theta_2, \ldots, \theta_n$,

$$\frac{m(\theta_2 \mid \theta_1) m(\theta_3 \mid \theta_2) \ldots m(\theta_n \mid \theta_{n-1}) m(\theta_1 \mid \theta_n)}{m(\theta_n \mid \theta_1) m(\theta_{n-1} \mid \theta_n) \ldots m(\theta_2 \mid \theta_3) m(\theta_1 \mid \theta_2)} = 1. \tag{20}$$

The Markov chain with transition probabilities $p(\theta' \mid \theta)$ given by the primitive experiment p satisfies the Kolmogorov criterion (20) since p is symmetric by assumption. Finally, for any positive termination strategy β, the effective choice rule $r(\theta' \mid \theta; p, \beta, \sigma_I)$ satisfies the Kolmogorov criterion too. This is because $r(\theta' \mid \theta; p, \beta, \sigma_I) = \frac{\beta_{\theta'} p(\theta'|\theta)}{\sum_{\tilde{\theta}} \beta_{\tilde{\theta}} p(\tilde{\theta}|\theta)}$, and when the expressions for $r(\theta' \mid \theta; p, \beta, \sigma_I)$ are substituted into (20), then the terms $\beta_{\theta'}$ and the denominators cancel out, and hence

$$\frac{r(\theta_2 \mid \theta_1; p, \beta\sigma_I) r(\theta_3 \mid \theta_2; p, \beta, \sigma_I) \ldots r(\theta_1 \mid \theta_n; p, \beta, \sigma_I)}{r(\theta_n \mid \theta_1; p, \beta, \sigma_I) r(\theta_{n-1} \mid \theta_n; p, \beta, \sigma_I) \ldots r(\theta_1 \mid \theta_2; p, \beta, \sigma_I)} = \frac{p(\theta_2 \mid \theta_1) p(\theta_3 \mid \theta_2) \ldots p(\theta_1 \mid \theta_n)}{p(\theta_n \mid \theta_1) p(\theta_{n-1} \mid \theta_n) \ldots p(\theta_1 \mid \theta_2)} = 1,$$

as needed. □

Proof of Proposition 3. Lemma 5 implies for all pairs $\theta, \theta' \in \Theta$:

$$r^*(\theta \mid \theta) r^*(\theta' \mid \theta) = r^*(\theta \mid \theta') r^*(\theta' \mid \theta').$$

By Lemma 2, we can substitute $r^*(\theta' \mid \theta) = \frac{\beta^*_{\theta'} p(\theta'|\theta)}{\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta}|\theta)}$, which gives

$$\frac{\beta^*_\theta \beta^*_{\theta'} p(\theta \mid \theta) p(\theta' \mid \theta)}{\left(\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta} \mid \theta)\right)^2} = \frac{\beta^*_\theta \beta^*_{\theta'} p(\theta \mid \theta') p(\theta' \mid \theta')}{\left(\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta} \mid \theta')\right)^2}.$$

Using symmetry of p we get

$$\frac{\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta} \mid \theta')}{\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta} \mid \theta)} = \frac{p^{1/2}(\theta' \mid \theta')}{p^{1/2}(\theta \mid \theta)}, \tag{21}$$

Equation (21), used below to prove the proposition, has an interesting interpretation on its own. It states that the decision rate $f_\theta = \sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta} \mid \theta)$ in each state θ is proportional to $p^{1/2}(\theta \mid \theta)$ and thus high in those states that are reliably identified by the primitive experiment p.

To compare $r^*(\theta_1 \mid \theta_2)$ and $r^*(\theta_2 \mid \theta_1)$, we write

$$\frac{r^*(\theta_1 \mid \theta_2)}{r^*(\theta_2 \mid \theta_1)} = \frac{\frac{\beta^*_{\theta_1} p(\theta_1|\theta_2)}{\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta}|\theta_2)}}{\frac{\beta^*_{\theta_2} p(\theta_2|\theta_1)}{\sum_{\tilde{\theta}} \beta^*_{\tilde{\theta}} p(\tilde{\theta}|\theta_1)}} = \frac{\frac{\beta^*_{\theta_1} p(\theta_1|\theta_2)}{p^{1/2}(\theta_2|\theta_2)}}{\frac{\beta^*_{\theta_2} p(\theta_2|\theta_1)}{p^{1/2}(\theta_1|\theta_1)}} = \frac{\beta^*_{\theta_1} p^{1/2}(\theta_1 \mid \theta_1)}{\beta^*_{\theta_2} p^{1/2}(\theta_2 \mid \theta_2)},$$

where we have used (21) in the second step, and symmetry of p in the last step. Define $\hat{\beta}_\theta = \beta^*_\theta p^{1/2}(\theta \mid \theta)$. We need to prove that if θ_1 is more distinct than θ_2, then $\hat{\beta}_{\theta_1} > \hat{\beta}_{\theta_2}$.

By (21), $(\hat{\beta}_\theta)_\theta$ satisfy the system of linear equations:

$$\sum_{\theta'} D_{\theta'\theta} \hat{\beta}_{\theta'} = 1 \text{ for all } \theta,$$

where $D_{\theta'\theta} = \frac{p(\theta'|\theta)}{p^{1/2}(\theta'|\theta') p^{1/2}(\theta|\theta)}$. We claim that if θ_1 is more distinct than θ_2, then $D_{\theta_3\theta_1} < D_{\theta_3\theta_2}$ for all $\theta_3 \neq \theta_1, \theta_2$. This follows from $p(\theta_3 \mid \theta_1) < p(\theta_3 \mid \theta_2)$ and from the symmetry of p:

$$p(\theta_1 \mid \theta_1) = 1 - p(\theta_2 \mid \theta_1) - \sum_{\theta_3 \neq \theta_1, \theta_2} p(\theta_3 \mid \theta_1) > 1 - p(\theta_1 \mid \theta_2) - \sum_{\theta_3 \neq \theta_1, \theta_2} p(\theta_3 \mid \theta_2) = p(\theta_2 \mid \theta_2),$$

and therefore,

$$D_{\theta_3 \theta_1} = \frac{p(\theta_1 \mid \theta_3)}{p^{1/2}(\theta_1 \mid \theta_1) p^{1/2}(\theta_3 \mid \theta_3)} < \frac{p(\theta_2 \mid \theta_3)}{p^{1/2}(\theta_2 \mid \theta_2) p^{1/2}(\theta_3 \mid \theta_3)} = D_{\theta_3 \theta_2}.$$

Thus,

$$D_{\theta_1 \theta_1} \hat{\beta}_{\theta_1} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_2} = 1 - \sum_{\theta_3 \neq \theta_1, \theta_2} D_{\theta_3 \theta_1} \hat{\beta}_{\theta_3} > 1 - \sum_{\theta_3 \neq \theta_1, \theta_2} D_{\theta_3 \theta_2} \hat{\beta}_{\theta_3} = D_{\theta_2 \theta_2} \hat{\beta}_{\theta_2} + D_{\theta_1 \theta_2} \hat{\beta}_{\theta_1}.$$

Using that $D_{\theta\theta} = 1$ and $D_{\theta\theta'} = D_{\theta'\theta}$, we have

$$\hat{\beta}_{\theta_1} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_2} > \hat{\beta}_{\theta_2} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_1}.$$

The assumption of sufficient precision of p and symmetry of p imply that $D_{\theta_2 \theta_1} < 1$, and thus $\hat{\beta}_{\theta_1} > \hat{\beta}_{\theta_2}$, as needed. □

A.4 Proofs for Section 6

Proof of Lemma 3. All rules feasible in $\mathcal{P}_{iia}$ are feasible in $\mathcal{R}(\mathcal{P}_{iia})$: $\mathcal{R}(\mathcal{P}_{iia}) \supset \mathcal{P}_{iia}$. This is immediate since when $\beta_a = 1$ for all $a \in A$, then $r(p, \beta, \sigma_I) = p$ for all $p \in \mathcal{P}_{iia}$.

It remains to show $\mathcal{R}(\mathcal{P}_{iia}) \subset \mathcal{P}_{iia}$. Consider $p(\gamma, \sigma) \in \mathcal{P}_{iia}$ constructed in the setting of Example 2 by the use of the generalized termination strategy $\gamma(m, x)$, and the action strategy $\sigma(m, x)$. Recall that $r(p(\gamma, \sigma), \beta, \hat{\sigma})$ is the choice rule constructed by repetitions of the rule $p(\gamma, \sigma)$ according to the termination strategy $\beta = (\beta_a)_{a \in A}$ and by applying the action strategy $\hat{\sigma} : A \longrightarrow A$ upon the termination. We need to show that there exists γ' and σ' such that $r(p(\gamma, \sigma), \beta, \hat{\sigma}) = p(\gamma', \sigma')$. This is indeed so when the termination probability $\gamma'(\mathfrak{t} \mid m, x) = \gamma(\mathfrak{t} \mid m, x)\beta_{\sigma(m,x)}$, the transition probability to the original memory state m_0 is $\gamma'(m_0 \mid m, x) = \gamma(m_0 \mid m, x) + \gamma(\mathfrak{t} \mid m, x)\left(1 - \beta_{\sigma(m,x)}\right)$, which is the sum of the probabilities that the original process γ transits to m_0 and that the decision process $r(p(\gamma, \sigma), \beta, \hat{\sigma})$ restarts after termination of $p(\gamma, \sigma)$. Additionally, for all $\tilde{m} \neq m_0$, $\gamma'(\tilde{m} \mid m, x) = \gamma(\tilde{m} \mid m, x)$. The above choice of γ' implies that the process $p(\gamma', \sigma')$ replicates the Markov process over the memory states under $r(p(\gamma, \sigma), \beta, \hat{\sigma})$. Finally, to replicate the choices upon terminations, we set the action strategy $\sigma'(m, x) = \hat{\sigma}(\sigma(m, x))$ for all (m, x). □

Proof of Lemma 4. Again, trivially, $\mathcal{R}(\mathcal{P}_{pf}) \supset \mathcal{P}_{pf}$, since $r(p, (1, \ldots, 1), \sigma_I) = p$ for all $p \in \mathcal{P}_{pf}$.

Additionally, $\mathcal{R}(\mathcal{P}_{pf}) \subset \mathcal{P}_{pf}$. This is indeed so because for any $\beta = (\beta_a)_{a \in A}$ and any $\hat{\sigma} : A \longrightarrow A$, $r(p(\gamma, \sigma), \beta, \hat{\sigma}) = p(\gamma', \sigma')$ where the termination probability $\gamma'(\mathfrak{t} \mid h, \theta) = \gamma(\mathfrak{t} \mid h, \theta)\beta_{\sigma(h)}$, the transition probability to the empty signal history $\emptyset$ is set to $\gamma'(\emptyset \mid h, \theta) = \gamma(\emptyset \mid h, \theta) + \gamma(\mathfrak{t} \mid$

$h,\theta)\left(1-\beta_{\sigma(h)}\right)$, and for all $\tilde{h} \neq \emptyset$, $\gamma'(\tilde{h} \mid h,\theta) = \gamma(\tilde{h} \mid h,\theta)$. Finally, the action strategy is set to $\sigma'(h) = \hat{\sigma}(\sigma(h))$ for all histories h. □

Proof of Proposition 4. We extend the definition of the effective experiment to the setting with discounting. Let

$$s_\delta(x \mid \theta; p, \beta) = \sum_t \sum_{\mathbf{x}^t : x_t = x} \delta^t \rho\left(\mathbf{x}^t \mid \theta; p, \beta\right),$$

where $\rho\left(\mathbf{x}^t \mid \theta; p, \beta\right)$ is the probability of the signal history $\mathbf{x}^t$ defined in (1). Thus, $s_\delta(x \mid \theta; p, \beta)$ is the discounted probability that the agent's last observed signal value is x. It satisfies the recursion:

$$s_\delta(x \mid \theta; p, \beta) = \beta_x p(x \mid \theta) + \delta \sum_{x' \in X} \left(1 - \beta_{x'} p\left(x' \mid \theta\right)\right) s_\delta(x \mid \theta; p, \beta), \tag{22}$$

where the first summand is the probability that the decision process terminates with x in the first round and the second summand is the discounted probability that the process terminates with x later. Solving (22) for s_δ gives

$$s_\delta(x \mid \theta; p, \beta) = \frac{\beta_x p(x \mid \theta)}{1 - \delta + \delta \sum_{x' \in X} \beta_{x'} p(x' \mid \theta)}.$$

The discounted repeated-cognition problem (13) is thus equivalent to

$$\max_{p \in \mathcal{P}, \beta \in B, \sigma \in S} \sum_{\theta \in \Theta, x \in X} \pi_\theta \frac{\beta_x p(x \mid \theta)}{1 - \delta + \delta \sum_{x' \in X} \beta_{x'} p(x' \mid \theta)} u(\sigma(x), \theta). \tag{23}$$

Consider x with an interior termination probability $\beta^*_x \in (0,1)$ and let $a = \sigma^*(x)$. The first-order condition of the problem (23) with respect to β_x is:

$$\begin{aligned}
\sum_{\theta \in \Theta} \pi_\theta \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta^*_x} u(a,\theta) - \delta \sum_{\theta \in \Theta, x' \in X} \pi_\theta s_\delta(x' \mid \theta; p^*, \beta^*) \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta^*_x} u(\sigma^*(x'), \theta) &= \\
\sum_{\theta \in \Theta} \pi_\theta \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta^*_x} u(a,\theta) - \delta \sum_{\theta \in \Theta, a' \in A} \pi_\theta r^*_\delta(a' \mid \theta; p^*, \beta^*, \sigma^*) \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta^*_x} u(a', \theta) &= 0,
\end{aligned}$$

where we have summed over all x' such that $\sigma^*(x') = a'$ in the second line. Multiplication by β^*_x and summation over all x such that $\sigma^*(x) = a$ gives (14). □

Proof of Proposition 5. Any (β, σ) that leads to a selection of only one action with certainty is dominated by the decision process that terminates after the first round and chooses an action equal to the observed signal value. Thus, both β^*_0 and β^*_1 are positive, and the action strategy is $\sigma^*(x) = x$ or $\sigma^*(x) = 1 - x$. Let us show that the action strategy σ^* must be the identity function σ_I.

Assume for contradiction that $\sigma^*(x) = 1 - x$. The payoff difference between the rule $r_\delta(p, \beta^*, \sigma^*)$

and the choice rule that always selects $a = 1$ must be positive, since the latter is dominated:

$$\begin{gathered}\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma^*) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma^*) - \pi_1 u_1 = \\ \pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_1 u_1 \geq \\ \pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) - \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma_I) > 0,\end{gathered}$$

where the first inequality follows from the fact that any discounted choice rule satisfies $\sum_a r_\delta(a \mid \theta; p, \beta, \sigma) \leq 1$. Similarly, the payoff difference between the rule $r_\delta(p, \beta^*, \sigma^*)$ and the rule that always selects $a = 0$ must be positive:

$$\begin{aligned}\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma^*) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma^*) - \pi_0 u_0 &= \\ \pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_0 u_0 &\geq \\ \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma_I) > 0.\end{aligned}$$

The last two inequalities imply:

$$\frac{r_\delta(0 \mid 1; p, \beta^*, \sigma_I)}{r_\delta(0 \mid 0; p, \beta^*, \sigma_I)} > \frac{\pi_0 u_0}{\pi_1 u_1} > \frac{r_\delta(1 \mid 1; p, \beta^*, \sigma_I)}{r_\delta(1 \mid 0; p, \beta^*, \sigma_I)}.$$

This establishes contradiction because as shown in the proof of Proposition 4, $r_\delta(x \mid \theta; p, \beta^*, \sigma_I) = s_\delta(x \mid \theta; p, \beta^*) = \frac{\beta^*_x p(x|\theta)}{1-\delta+\delta \sum_{x'} \beta^*_{x'} p(x'|\theta)}$, and thus

$$\frac{r_\delta(1 \mid 1; p, \beta^*, \sigma_I) r_\delta(0 \mid 0; p, \beta^*, \sigma_I)}{r_\delta(0 \mid 1; p, \beta^*, \sigma_I) r_\delta(1 \mid 0; p, \beta^*, \sigma_I)} = \frac{p(1 \mid 1) p(0 \mid 0)}{p(0 \mid 1) p(1 \mid 0)} > 1.$$

Further, it must hold that $\beta^*_0 = 1$ or $\beta^*_1 = 1$. Otherwise, if both $\beta^*_0 < 1$ and $\beta^*_1 < 1$, then the agent can increase both β^*_x by a same factor. This preserves the conditional action distribution in each state θ and increases the decision rates in both states, and thus it is a profitable deviation.

Additionally, it must be that $\beta^*_1 = 1$: Using the expressions for $s_\delta(\theta \mid \theta; p, \beta) = r_\delta(\theta \mid \theta; p, \beta, \sigma_I)$, the payoff for σ_I and $(\beta_0, \beta_1) = (\beta, 1)$ is

$$\pi_0 u_0 \frac{\beta\alpha}{1 - \delta + \delta(\beta\alpha + 1 - \alpha)} + \pi_1 u_1 \frac{\alpha}{1 - \delta + \delta(\alpha + \beta(1 - \alpha))}, \tag{24}$$

and payoff for σ_I and $(\beta_0, \beta_1) = (1, \beta)$ is

$$\pi_0 u_0 \frac{\alpha}{1 - \delta + \delta(\alpha + \beta(1 - \alpha))} + \pi_1 u_1 \frac{\beta\alpha}{1 - \delta + \delta(\beta\alpha + 1 - \alpha)}, \tag{25}$$

The assumptions that $\pi_1 u_1 > \pi_0 u_0$ and that $\alpha > 1/2$ imply that, for any $\beta \in (0, 1)$, (24) exceeds (25), as needed.

It therefore remains to find $\beta^*_0 \in (0, 1]$. If the optimal value is interior, then it satisfies (14) with $a = 0$:

$$\pi_0 u_0 r_\delta(0 \mid 0, p, \beta^*, \sigma_I) = \delta \left(\pi_0 u_0 r_\delta^2(0 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma_I) r_\delta(0 \mid 1; p, \beta^*, \sigma_I)\right).$$

After the substitution of $r_\delta(x \mid \theta; p, \beta, \sigma_I) = \frac{\beta_x p(x|\theta)}{1-\delta+\delta\sum_{x'}\beta_{x'}p(x'|\theta)}$, this condition simplifies into a quadratic equation for β_0^*. When $\delta < \frac{1}{\alpha+(1-\alpha)R}$, then this condition does not have an interior solution and the derivative of the value (24) with respect to β_0 at $\beta_0 = 1$ is positive. Thus, in this case, the unique β_0^* satisfying the first-order condition is $\beta_0^* = 1$.

When $\delta > \frac{1}{\alpha+(1-\alpha)R}$, then the condition has an interior solution and the derivative of the value (24) with respect to β_0 at $\beta_0 = 1$ is negative. Thus, for this range of parameters, the unique β_0^* satisfying the first-order condition is the interior value that solves the quadratic equation, solution of which decreases in δ. □

A.5 Proofs for Section 7

Proof of Corollary 7. Note that the choice rule r^* characterized in (10) is homogeneous with degree zero with respect to u_0 and u_1. Thus, without loss of generality, we can set $u_0 = 1$. Bayes rule implies

$$\begin{aligned}
\Pr_{r^*}(\theta = 1 \mid a = 1) &= r^*(1 \mid 1)\frac{\pi_1}{\pi_1 r^*(1 \mid 1) + (1-\pi_1)(1 - r^*(0 \mid 0))},\\
\Pr_{r^*}(\theta = 1 \mid a = 0) &= (1 - r^*(1 \mid 1))\frac{\pi_1}{\pi_1(1 - r^*(1 \mid 1)) + (1-\pi_1) r^*(0 \mid 0)}.
\end{aligned}$$

We substitute the expressions for $r^*(0 \mid 0)$ and $r^*(1 \mid 1)$ from (10) and take derivatives with respect to π_1 to get

$$\begin{aligned}
\frac{\partial}{\partial \pi_1}\Pr_{r^*}(\theta = 1 \mid a = 1) &= \frac{\bar{d}\sqrt{\frac{\bar{d}\pi_1 u_1}{1-\pi_1}}}{2\left(\bar{d}\pi_1 + (1-\pi_1)\sqrt{\frac{\bar{d}\pi_1 u_1}{1-\pi_1}}\right)^2},\\
\frac{\partial}{\partial \pi_1}\Pr_{r^*}(\theta = 1 \mid a = 0) &= \frac{\sqrt{\frac{\bar{d}\pi_1 u_1}{1-\pi_1}}}{2\left(\pi_1 + (1-\pi_1)\sqrt{\frac{\bar{d}\pi_1 u_1}{1-\pi_1}}\right)^2},
\end{aligned}$$

where both derivatives are positive, as needed. □

References

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica 17*(3), 213–244.

Basu, P. and K. Chatterjee (2015). On interim rationality, belief formation and learning in decision problems with bounded memory.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics 127*(3), 1243–1285.

Calvert, R. L. (1985). The value of biased information: A rational choice model of political advice. *The Journal of Politics 47*(2), 530–555.

Che, Y.-K. and K. Mierendorff (2016). Optimal sequential decision with limited attention. *unpublished, Columbia University*.

Compte, O. and A. Postlewaite (2012). Belief formation. PIER Working Paper 12-027, University of Pennsylvania.

Dow, J. (1991). Search decisions with limited memory. *The Review of Economic Studies 58*(1), 1–14.

Fryer, R. G., P. Harms, and M. O. Jackson (2016). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization.

Fudenberg, D., P. Strack, and T. Strzalecki (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*.

Gabaix, X. and D. Laibson (2017). Myopia and discounting. Technical report, National Bureau of Economic Research.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of Political Economy 114*(2), 280–316.

Hébert, B. and M. Woodford (2016). Rational inattention with sequential information sampling. *unpublished, Stanford University*.

Hellmann, M. and T. M. Cover (1970). Learning with finite memory. *Annals of Mathematical Statistics 41*, 765–782.

Itti, L. (2007). Visual salience. *Scholarpedia 2*(9), 3327.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory 123*(2), 81–104.

Jehiel, P. (2018, June). Investment strategy and selection bias: An equilibrium perspective on overoptimism. *American Economic Review 108*(6), 1582–97.

Johnson, D. D., D. T. Blumstein, J. H. Fowler, and M. G. Haselton (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology & Evolution 28*(8), 474–481.

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Khaw, M. W., Z. Li, and M. Woodford (2017). Risk aversion as a perceptual bias. Technical report, National Bureau of Economic Research.

Leung, T. K. (2017). Limited cognitive ability and selective information processing. *unpublished, Toulouse School of Economics*.

Lieder, F., M. Hsu, and T. L. Griffiths (2014). The high availability of extreme events serves resource-rational decision-making. In *Proceedings of the Cognitive Science Society*, Volume 36.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* Number 8. Oxford University Press on Demand.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *The American Economic Review 105*(1), 272–298.

Meyer, M. A. (1991). Learning from coarse information: Biased contests and career profiles. *The Review of Economic Studies 58*(1), 15–41.

Morris, S. and P. Strack (2017). The wald problem and the equivalence of sequential sampling and static information costs.

Netzer, N. (2009). Evolution of time preferences and attitudes toward risk. *The American Economic Review 99*(3), 937–955.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology 2*(2), 175.

Oswald, M. E. and S. Grosjean (2004). Confirmation bias. *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory 79.*

Piccione, M. and A. Rubinstein (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior 20*(1), 3–24.

Rabin, M. and J. L. Schrag (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics 114*(1), 37–82.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review 85*(2), 59.

Ratcliff, R. and G. McKoon (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation 20*(4), 873–922.

Rayo, L. and G. S. Becker (2007). Evolutionary efficiency and happiness. *Journal of Political Economy 115*(2), 302–337.

Robson, A. J. (2001). The biological basis of economic behavior. *Journal of Economic Literature 39*(1), 11–33.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Steiner, J., C. Stewart, and F. Matějka (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica 85*(2), 521–553.

Suen, W. (2004). The self-perpetuation of biased beliefs. *The Economic Journal 114*(495), 377–396.

Swensson, R. G. (1972). The elusive tradeoff: Speed vs accuracy in visual discrimination tasks. *Perception & Psychophysics 12*(1), 16–32.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics 16*(2), 117–186.

Wilson, A. (2014). Bounded memory and biases in information processing. *Econometrica 82*(6), 2257–2294.

Zhong, W. (2017). Optimal dynamic information acquisition. *unpublished, Columbia University*.