Cost-Benefit Analysis in Reasoning*

Larbi Alaoui[†]
UPF and BSE

Antonio Penta[‡]
ICREA, UPF and BSE

December 1, 2021

Abstract

An individual's decision to reason about a problem may involve a tradeoff between cognitive costs and a notion of value. This paper analyzes the primitive properties that must hold for the decision to stop thinking to be represented by a cost-benefit analysis. We then provide additional properties that give more structure to the value of reasoning function. We show how our model applies to a variety of settings, including contexts involving R&D applications, response time, and strategic reasoning. Our model can also be used to understand patterns of behavior for which the cost-benefit approach does not seem to hold.

Keywords: cognition and incentives – choice theory – reasoning – fact-free learning

JEL Codes: D01; D03; D80; D83.

*Alaoui gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation under projects ECO2014-56154-P, PGC2018-098949- B-I00, grant RYC-2016-21127 (Alaoui), as well as through as through the Severo Ochoa Program for Centers of Excellence in R&D, grants SEV-2015-0563 and CEX2019-000915-S. Penta acknowledges the financial support of the ERC Starting Grant #759424. Early versions of this paper have been presented at seminars at Princeton, MIT-Harvard, Northwestern, Columbia, Minnesota, Carnegie-Mellon, UCLA, Boston University, Bocconi, UPF, and many conferences. We are grateful to those audiences for the invaluable comments. Special thanks go to Jose Apesteguia, Andrew Caplin, Gabriel Carroll, Sylvain Chassang, Bart Lipman, Fabio Maccheroni, George Mailath, Paola Manzini, Marco Mariotti, Laurent Mathevet, Efe Ok, Pietro Ortoleva, Andy Postlewaite, Aldo Rustichini, Bill Sandholm, Alvaro Sandroni, Marciano Siniscalchi, the editor Ali Hortaçu, and two anonymous referees. Edited by Ali Hortaçu.

[†]Universitat Pompeu Fabra and Barcelona School of Economics. E-mail: larbi.alaoui@upf.edu

[‡]ICREA, Universitat Pompeu Fabra and Barcelona School of Economics. E-mail: antonio.penta@upf.edu

1 Introduction

Making a difficult choice or approaching a hard problem requires thought and introspection. In proving a theorem, all of logic is at our disposal, and yet we rarely reach the solution instantly or are even certain of the truth of the proposition – instead, we reason in steps, pushing the frontier of our understanding, sometimes in leaps of insight and sometimes in small increments, and sometimes never arriving to the solution. Similarly, choosing between complex options often leads people to resort to a deliberation process that relies on heuristics. The reasoning process does not involve receiving new data, but, because of cognitive limitations and failures of logical omniscience, fact-free learning can occur (cf. Aragones, Gilboa, Postlewaite and Schmeidler (2005)). Furthermore, individuals often stop thinking before achieving full understanding. This need not imply they have reached an absolute cognitive limit. An increase in the academic or financial value of proving a theorem may well provide the researcher with the incentives to reason further. Similarly, higher stakes for making a choice could lead the consumer to be more deliberative. This suggests that individuals trade off the cognitive cost against some possibly vague notion that reasoning is worthwhile. Hence, even though fact-free learning (or 'reasoning') spans a wide range of contexts, including computational complexity (e.g., Aragones et al. (2005) and Gabaix and Laibson (2005)), unawareness (e.g., Schipper (2015)), bounded rationality in games (see Camerer (2003) and Crawford, Costa-Gomes and Iriberri (2013)) and sequential heuristics in choice (e.g., Manzini and Mariotti (2007)), the underlying features are the same. Reasoning is a stepwise process, it is cognitively costly, and individuals may find it worthwhile.

The first instinct to an economist is then to cast the decision to reason further as a cost-benefit tradeoff. But this approach can be problematic, and is often cautioned against in psychology and cognitive science. Moreover, the subjective value that an individual attaches to reasoning is not well-defined: it is not clear how an individual should assess the benefit of reasoning further when he has no sense of what it might bring, especially when he has only a partial understanding of the problem. It is also not obvious how to capture a researcher's outlook over what thinking more about a proof will lead to. Understanding the relation between incentives and cognition is important for many economic settings, but little is known about when a cost-benefit analysis is justified, about the form it should take, or about which properties of the reasoning process must fail when it does not hold.

These are entirely new questions. To address them, we develop a framework to model the reasoning process and identify the core properties that are necessary and sufficient for reasoning to be captured by a cost-benefit analysis. We find that they are weak and intuitive, suggesting that there is indeed justification for using cost-benefit models for a large class of problems. But our conditions also identify the limits of the approach: in contexts for which these properties do not hold, reasoning would not follow a cost-benefit criterion. We also provide conditions that give structure to the value of reasoning function, including a 'value of information' representation, often assumed in the literature (e.g., Caplin and Dean (2015)), as well as a novel 'maximum gain' representation, which is particularly useful for applications.

We model an iterated reasoning process as a sequence of mental states. A mental state encodes both the agent's current understanding of the problem at hand, and his outlook towards what performing an additional step of reasoning may bring. Specifically, each mental state is summarized by two objects: an action that he finds best, given his current understanding, and a preference relation, which describes the agent's understanding at that stage, what he expects to learn from reasoning further, and his choice of whether to stop thinking or not. We then formalize properties of the reasoning process as axioms over these preferences, and use them to obtain representations of the subjective costs and benefits, which are not directly observable to the experimenter.

The key properties in our model can be informally summarized in a simple way: the agent reasons only if it is relevant to his choice; at any mental state, when reasoning has no instrumental value, the agent's reluctance to think is identical for two problems that only differ by a constant payoff; and he should not have a strict preference for committing to ignore what he might learn. Our characterization in terms of these properties therefore identifies the testable implications of adopting a cost-benefit approach to reasoning, thereby providing a precise demarcation of its domain of applicability. Since these properties are all intuitive, our results also show that a cost-benefit approach can be a useful 'as if' model for various reasoning processes, thereby justifying the applications of basic economics concepts to these novel domains. We also discuss additional properties and what they imply for the shape of the value of reasoning function. Under some conditions, the value function takes a 'maximum gain' shape – it is as if the value attached to reasoning corresponds to the highest payoff improvement the agent could obtain, relative to his current most-preferred action. Under alternative conditions, it is as if the agent's outlook is not as extreme, and more measured in what reasoning might bring. For instance, the value function can take a 'value of information' form, for which it is as if, at each mental state, the agent has subjective beliefs over the outcome of the next step of reasoning, and the value equals the expected gain of switching from the current action to the future best response. We discuss how these conditions, which relate current understanding and outlook towards future reasoning, might seem plausible in some settings but not in others.

The advantage of a cost-benefit representation is that it allows for comparative statics exercises on the depth of reasoning. But for these to be meaningful, there must be a way of shifting the value without changing the cost or the process itself. We therefore introduce a notion of 'cognitive equivalence class', which groups together problems that are equally difficult and approached essentially in the same way. Our representation theorems imply that the costs are uniquely pinned down within each class. This ensures that cognitively equivalent problems only differ in the value of reasoning they entail, thereby allowing for useful comparative statics on the depth of reasoning. We also obtain a strong uniqueness result in our representations, for both the cost and value functions.

This framework can be applied to diverse theoretical and empirical settings in which cognition is costly. We first illustrate how the model can be used in an R&D domain, and how the dataset we require to test our axioms can be elicited. We also propose an experiment designed to elicit such preferences in the lab. In applications in which reasoning is mainly introspective, such elicitation methods are harder to implement, and hence a strict revealed preference interpretation is more problematic. Nonetheless, the preference relations in our framework remain clear, conceptually, and provide a language to identify the cognitive building blocks of different modes of reasoning. Moreover, even in these settings the model's joint implications can be tested. To illustrate the point, we first consider an extension to a model of endogenous level-k reasoning (Alaoui and Penta, 2016a), which

$$u = \begin{array}{c|c|c|c|} & \omega_1 & \omega_2 & \omega_3 \\ \hline a_1 & 1 & 0 & 0 \\ \hline a_2 & 0 & 1 & 0 \\ \hline a_3 & 0 & 0 & 1 \\ \hline \end{array}$$

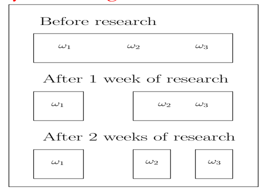

Figure 1: R&D Example

has strong experimental support and is predictive of well-known behavioral anomalies. Then, we provide an extension of the model to connect our notion of depth of reasoning with measures of response time, and discuss recent experimental evidence by Alos-Ferrer and Buckenmaier (2019) in support of the model. Finally, we discuss how the model can be used for estimation.

Our analysis is useful not only for understanding the scope of the cost-benefit approach, but also for exploring its limits. In particular, there is a growing literature in economics on contexts where these limits are apparent. These include settings in which individuals display 'thinking aversion', and in tasks for which higher financial rewards are sometimes detrimental to performance (e.g. Gneezy and Rustichini (2000) or Camerer and Hogarth (1999)). A number of mechanisms have been proposed in the psychology literature to explain this observed behavior, such as different forms of choking under pressure, anxiety, cognitive overload or distractions. These mechanisms are often intertwined and difficult to disentangle theoretically and empirically. Our model allows a clear separation of these theories, which we show by mapping each mechanism to specific violations of our axioms. Identifying these properties can then provide a basis for both theoretical and empirical work, as we discuss below.

2 The Model of Reasoning

In this section we introduce the main elements of our framework, which we first illustrate in the context of a simple example.

2.1 An R&D Example

Consider the problem of a researcher working in a firm's R&D department, who needs to identify which of three products $a \in A = \{a_1, a_2, a_3\}$ is best to take to production. The correct answer depends on some state of the world, $\omega \in \Omega = \{\omega_1, \omega_2, \omega_3\}$, but he does not know which state is the true one. For simplicity, let's suppose that the researcher obtains util payoff 1 if he chooses the right product at the corresponding state (a_i is correct if the true state is ω_i), and 0 otherwise, as described by the payoff matrix u on the left in Figure 1.

At the end of every week the researcher has spent on the problem, he prepares a report which contains the action (product) which he currently considers to be best, as well as a recommendation to stop the investigation and commit to a product, or postpone the decision and conduct more research. If research continues, the researcher will prepare a new report at the end of the next week, and so on, until a decision to stop is made. The sequence of experiments to be done has been decided in advance at board level, so the only question is whether to continue to run the experiments or not.

Suppose that the researcher's prior beliefs over states ω_1, ω_2, and ω_3 are, respectively, $5/8, 1/4$, and $1/8$, and suppose that his investigation strategy consists of first checking whether ω_1 is the true state of the world or not. If he discards it then he can test ω_2, and either confirms that the true state is ω_2, or that it must be ω_3.

Now, suppose that the true state of the world is ω_3. In this situation, if the researcher were not to investigate at all, and make the product recommendation purely based on his prior beliefs, then he would choose a_1. If he were to postpone, and at the next step discard ω_1, then he could again decide whether to investigate further or not. If he is Bayesian (which we will not require for the general model), then his beliefs would be updated to $2/3$ for ω_2 and $1/3$ for ω_3. If he stops then he picks a_2, but if he decides to investigate further then he discards ω_2, and is certain that the state is ω_3. He therefore chooses a_3. He has of course the option to research more, but nothing else would be learned at that point, and he would remain with the certainty that the state is ω_3.

As a benchmark, if the agent is Bayesian and conducts a standard, fully sophisticated cost-benefit analysis, then he would not stop at step $k-1$ and continue reasoning if $W^*(u, k-1) > E_\omega u(a^{k-1}, \omega) + c(k)$, where $W^*(u, k-1)$ is the value function (in a Bellman sense) of continuing reasoning at the end of week $k-1$, $c(k)$ is the cost of the next period of investigation, and $E_\omega u(a^{k-1}, \omega)$ is the expected utility of choosing the best action, a^{k-1}. Note that, in this fully sophisticated model, the value function $W^*(u, k-1)$ includes the entire optimal planning for the full dynamic process, conditional on each realization of the outcome of future research which is still unknown to the researcher. Hence, it includes not only the value of choosing the optimal action based on the information at each future node, but also the option value from stopping or continuing at every contingency.

As this is a highly complex problem for an agent or a firm to face, or for an applied economist to use, simpler kinds of stopping rules may be more apt in some settings, or more amenable to tractable analysis. For instance, the agent may still be Bayesian, but be *myopic* in the sense that his decision to stop or continue is only based on what he thinks he might learn if he continues, without fully calculating the costs and option values of further postponing his choice in the future.[1] In that case, given the action a^{k-1} the agent regards as best at the end of step $k-1$, his *value of reasoning* would be as follows:

$$W^1(u, k-1) = \sum_{\mu \in \Delta(\Omega)} p^k(\mu) \sum_{\omega} \mu(\omega) \cdot u(a^*(\mu), \omega),$$

where μ are beliefs over ω that the agent may hold after conducting more research, p^k is a probability distribution over such μ, which represents the researcher's beliefs on what he might learn, and $a^*(\mu)$ is the optimal action given μ. In that case, he would continue if

$$\underbrace{\sum_{\mu \in \Delta(\Omega)} p^k(\mu) \sum_{\omega} \mu(\omega) u(a^*(\mu), \omega)}_{W^1(u,k-1)} > \underbrace{\sum_{\omega} \mu^{k-1}(\omega) u(a^{k-1}, \omega)}_{E_\omega u(a^{k-1},\omega)} + c(k),$$

[1] We do not discuss here the fully sophisticated agent, which significantly complicates the analysis. However, that discussion is provided in Alaoui and Penta (2021), which extends the model presented here, and provides an axiomatic characterization of the relevant representation.

(where, by Bayesian consistency, $\mu^{k-1}(\omega) = \sum_{\mu \in \Delta(\Omega)} p^k)$ or, in terms of *net benefit*,

$$V(u,k) = \underbrace{\sum_{\mu \in \Delta(\Omega)} p^k(\mu) \sum_{\omega} \mu(\omega)[u(a^*(\mu),\omega) - u(a^{k-1},\omega)]}_{\substack{W^1\left(u-u^{k-1},k-1\right), \\ \text{where } u^{k-1}(a,\omega) := u(a^{k-1},\omega) \text{ for all } (a,\omega) \in A \times \Omega.}} > c(k).$$

That is, the agent continues reasoning if the expected gain of switching from a^{k-1} to the optimal response to the future signals is larger than the cost. We note that both the fully sophisticated and myopic representations share some common properties. For instance, in both cases: (i) the (net) benefit of continuing is never negative; (ii) if a constant payoff of $t \in \mathbb{R}$ were added to each cell of the matrix of outcomes (that is, if payoffs were transformed from $u(a,\omega)$ to $u'(a,\omega) = u(a,\omega) + t$ for each (a,ω)), then at no point of the investigation would the researcher's decision to continue researching be affected; (iii) if all payoffs where multiplied by a constant $\alpha > 1$, then the researcher's incentives to postpone making a choice and continue researching would increase at every k; (iv) if instead we let such an α approach 0, then the researcher would choose not to reason for progressively lower costs, and never strictly prefer to continue researching if payoffs u are constant in a.

A different researcher, however, need not base his decisions on either of the above criteria, but still have some notion of cost-benefit trade-off in mind. For instance, for a particularly zealous researcher who constantly challenges his current understanding, the (net) value of reasoning could be driven by the maximum utility gain that he could obtain by continuing his research, as in the following representation:

$$V(u,k) = \max_{\omega \in \Omega} u(a^*(\omega),\omega) - u(a^{k-1},\omega),$$

where $a^*(\omega)$ is the optimal action at ω. We will characterize the properties of the reasoning process under which the agent's decision to stop or continue takes the form of a general cost-benefit analysis, a myopic value of information, the maximum gain representation, and others. We will also explore what these representations have in common (for instance, they all satisfy properties (i) through (iv) above) and the ways in which they differ.

We obtain these characterizations from the agent's primitive preferences over continuing reasoning or not, for various payoffs. To elicit these preferences, we assume that we can stop the agent at the end of any week of research, right after his report, and ask any of the following three kinds of questions. Each will correspond to a specific part of the cost-benefit representation, which will be formalized in Definition 2.

Q.0: Given the research he has conducted (i.e., the understanding that he has reached) and the action recommendation he has just made, we can ask the researcher whether he prefers being awarded according to his original payoff matrix u, or another v, without the possibility of conducting more research. These comparisons will be represented by function W^0 in Def. 2.

Q.1: We can ask, under the condition that he has to perform at least one extra week of research, and without changing the firm's research plan, whether he prefers being rewarded according to function u or v. (Note that we do not force him to stop research after one week, and his choice of when to stop can depend on whether rewards are determined by u or v.) As in the example above, these comparisons will be represented by a function denoted W^1 in Def. 2.

Q.0-1: Lastly, we can ask him whether he prefers to stop now and be rewarded according to u, or continue research, still without changing the research plan, and be rewarded according to a payoff matrix $v = u + t$, where $t \in \mathbb{R}$ is some constant compensation that may be added to u. These comparisons will be captured by cost function c and value of reasoning V in Def. 2. As in the example above, the latter will take the form $V(u,k) = W^1(u - u^{k-1}, k-1)$.

Question Q.0-1 are the ones that we are ultimately interested in, in that they capture the decision to reason further or not in any given decision problem. These are the comparisons that are typically most relevant, and the ones for which the very notion of *value of reasoning* is defined (cf. the example above). To obtain such representations, within which comparative statics across u are well-defined, we require an extension of our datasets, hence the additional questions. Nonetheless, we need not observe the entire extension of this dataset to obtain our representation theorems, which we view as an advantage of this approach.

2.2 Model

We explain next how we formalize the ideas illustrated above. An agent chooses from a finite set A of acts, which map states $\omega \in \Omega$ to utils, which for simplicity we assume have already been elicited, and are therefore observed as well. That is, for any act $a \in A$, there is a function $u(a,\cdot) : \Omega \to \mathbb{R}$ from the states of the world to final util payoffs. The menu A of acts can thus be represented by the overall function $u : A \times \Omega \to \mathbb{R}$. We let $\mathcal{U} := \mathbb{R}^{|A\times\Omega|}$ denote the set of menu of acts. Elements of $A \times \Omega$ are referred to as *outcomes*.

In practice, the agent's reasoning may be affected by features other than the payoff function, such as: the true state of nature (in the earlier example, the optimal action after the first week is a_1 if the true state is ω_1, and a_2 otherwise); the complexity of the problem (holding constant the payoff consequences, understanding which product is optimal may require complex research, or just some reading); or other features of the context (e.g., the researcher's decisions may change if he worked for a different firm); and so on. To capture this, we let an *environment* E denote a full list of all the elements which may affect the agent's reasoning (including the true state), and let $\mathcal{E}$ denote the set of possible environments. The agent's overall *choice problem* is thus defined as a pair $(u, E) \in \mathcal{U} \times \mathcal{E}$.

Notation: For any $u, v \in \mathcal{U}$ and $\alpha \in [0,1]$, we denote by $\alpha u + (1-\alpha) v \in \mathcal{U}$ the utility function that pays $\alpha u(a,\omega) + (1-\alpha) v(a,\omega)$ for each $(a,\omega) \in A \times \Omega$. It will sometimes be useful to think of utility functions as vectors in $\mathbb{R}^{|A\times\Omega|}$, and write functions that pay $t \in \mathbb{R}$ for all outcomes as $t \cdot \mathbf{1}$, where $\mathbf{1}$ denotes the unit vector. For $t \in \mathbb{R}$, we will also write $u + t$ to mean $u + t \cdot \mathbf{1}$. We adopt the following conventions for vectors: for any $x, y \in \mathbb{R}^n$, we let $x > y$ denote the case in which x is weakly larger than y for each component, but not $x = y$; $x \geq y$ means $x > y$ or $x = y$; we let $x >> y$ denote strict inequality for each component. For any set X, we let $\Delta(X)$ denote the set of simple (i.e., finite support) probability measures on X. Finally, for functions $f : X \times Y \to \mathbb{R}$, we write $f(\cdot, y)$ to denote the function $g : X \to \mathbb{R}$ s.t. $g(x) = f(x,y)$ for all $x \in X$.

2.2.1 Mental States, Actions, and Preferences

Each step of reasoning is described by a *mental state* $s \in S$, which consists of both the action that the agent finds best at s (i.e., if he were to stop), denoted a^s, and a preference relation $\succsim_s$ on $\mathcal{U} \times \{0,1\}$, with asymmetric and symmetric parts $\succ_s$ and $\sim_s$, respectively.

The projection $\succsim_s$ on $\mathcal{U} \times \{0\}$, denoted by $\succsim_s^0$, describes the agent's *current understanding*: his ranking of payoff matrices u and v given the understanding he has attained at s, if he were to choose now, without thinking further. These preferences correspond to the answers to questions Q.0 in the leading example.

The projection $\succsim_s$ on $\mathcal{U} \times \{1\}$, denoted by $\succsim_s^1$, describes instead the agent's *outlook over his* future *understanding*: suppose that, in the reasoning process about choice problem (u, E), the agent has reached mental state s; then $\succsim_s^1$ elicits his comparison between u and v if he were to think further, just as in questions Q.1 in the leading example.

Finally, for the 0-1 comparisons, our main focus is on whether the agent prefers to stop thinking at s (i.e., $(u,0) \succsim_s (u,1)$) or to continue (i.e., $(u,1) \succsim_s (u,0)$), where 'continue' only means to postpone the choice, with no commitment on when. However, as explained with questions Q.0-1 in the leading example, we also elicit the agent's comparisons between stopping at s and be rewarded according to u, or continuing and receive an extra reward $t \in \mathbb{R}$ (e.g., $(u,0) \succsim_s (u+t,1)$ or $(u+t,1) \succsim_s (u,0)$).

We maintain throughout that $\succsim_s$ compares all pairs $(u,x),(v,y) \in \mathcal{U} \times \{0,1\}$ s.t. either (i) $y = x$, or (ii) $x = 0$, $y = 1$ and $v = u + t$ for some $t \in \mathbb{R}$ (that is, $(u,x) \succsim_s (v,y)$ or $(v,y) \succsim_s (u,x)$), and that it admits a complete and transitive extension to $\mathcal{U} \times \{0,1\}$.

Our dataset includes both a^s and $\succsim_s$, and it can be easily collected in settings such as the R&D case of our leading example. Our model, however, also applies to other settings, such as trying to prove a theorem or other forms of introspective reasoning. In those cases the elicitation may be less straightforward (see Section 4.2), but the conceptual exercise remains simple and serves to understand the primitive properties of reasoning underlying the representations we will discuss.

2.2.2 Reasoning, Cognitive Equivalence, and Cost-Benefit Representation

A *reasoning process* is a sequence of mental states $(s^k)_{k\in\mathbb{N}}$. For each choice problem $(u,E) \in \mathcal{U} \times \mathcal{E}$, we let $(s^k_{(u,E)})_{k\in\mathbb{N}}$ denote the reasoning process it induces, and let $S(u,E) := \bigcup_{k\in\mathbb{N}}\{s^k_{(u,E)}\}$. The dependence of the reasoning process on both u and E formalizes the idea that an agent may reason differently in different problems, or in the same problem at a different realized state (which, as discussed above, is embedded in E).[2]

Different choice problems, however, may be approached in very different ways, and it would not be meaningful for instance to compare the depth of reasoning of the researcher in the leading example across environments in which he has access to different technologies, or for wildly different payoff

[2] In case of stochastic reasoning, the realization of the stochastic process is embedded in E, and hence the $(s^k_{(u,E)})_{k\in\mathbb{N}}$ is an ex-post realization. Since, at any mental state, the agent does not know the future states, and his beliefs over future steps are not required to be correct, the distinction does not matter. An ex-ante approach would be useful for comparisons across individuals, or when imposing correct beliefs. The fact that the path of reasoning also involves mental states that follow the one at which, according to our representation, the agent would stop reasoning, enables us to accommodate cases in which individuals' reasoning do not satisfy our properties (see, e.g., Section 5.2, in which we discuss violations of cost-benefit).

functions: in these cases, the agent's research strategy may be entirely different, and entail very different costs, incentives, and paths of reasoning. To formalize these ideas, we introduce the notion of cognitive equivalence: In essence, two problems are cognitively equivalent if the agent approaches them in exactly the same way.

Definition 1 (Cognitive Equivalence)**.** *Two mental states $s, \hat{s} \in S$ are* cognitive equivalent (c.e.) *if $a^s = a^{\hat{s}}$ and $\succsim_s = \succsim_{\hat{s}}$. Two choice problems (u, E) and $(\hat{u}, \hat{E})$ are* cognitively equivalent *if they induce sequences of pairwise c.e. states, s^k and $\hat{s}^k$ for each k.*

We let $\mathcal{C}$ denote the *cognitive partition* on $\mathcal{U} \times \mathcal{E}$ induced by the cognitive equivalence relation, and let $C \in \mathcal{C}$ denote a generic cognitive equivalence class. We also let $C(u, E)$ denote the equivalence class which consists of all problems that are cognitively equivalent to (u, E). Problems (u, E) and $(\hat{u}, \hat{E}) \in C(u, E)$ induce the same path of reasoning, in that they induce sequences $(s^k)_{k \in \mathbb{N}}$ and $(\hat{s}^k)_{k \in \mathbb{N}}$ that correspond to the same action a^k and preferences $\succsim_k$ for each k. We will use $(a^s, \succsim_s)$ to refer to the action associated with an arbitrary s, and $(a^k, \succsim_k)$ for the action and preference relation at the k-th state s^k along a specific path.

In Section 3 we will assume that any two choice problems are c.e. if they only differ in the payoff function, with one being a positive affine transformation of the other. Intuitively, that assumption will require that, if a constant is added to all payoffs, or if all payoffs are multiplied by the same positive scalar, holding the environment constant, it will not change the way the agent reasons about the problem. In all representation theorems that follow, cognitively equivalent problems have the same associated cost of reasoning at any step k. In other words, in our representations, problems that are cognitively equivalent not only share the same path of reasoning, but also the same costs. This will be formalized in Definition 2.

In the following, an important role will be played by the collection of 'payoff differences'

$$\Big(u\,(a, \omega) - u(a^k, \omega)\Big)_{(a,\omega) \in A \times \Omega}, \tag{1}$$

which describes the potential gains (losses, if the entry is negative), given payoff function u, from switching from the current choice a^k to some other a, at every ω. For instance, for the u in our leading example, and for $a^k = a_1$, we have:

$$(u\,(a, \omega) - u\,(a_1, \omega))_{(a,\omega) \in A \times \Omega} = \begin{array}{|c|c|c|c|} \hline & \omega_1 & \omega_2 & \omega_3 \\ \hline a_1 & 0 & 0 & 0 \\ \hline a_2 & \text{-1} & 1 & 0 \\ \hline a_3 & \text{-1} & 0 & 1 \\ \hline \end{array}$$

Note that, for any $\hat{a}$, $(u\,(a, \omega) - u\,(\hat{a}, \omega))_{(a,\omega) \in A \times \Omega} = \mathbf{0}$ if and only if u is constant in a. Finally, for any $u \in \mathcal{U}$, we define $u^k : A \times \Omega \to \mathbb{R}$ such that $u^k(a, \omega) = u(a^k, \omega)$ for all $(a, \omega) \in A \times \Omega$. (With this notation, the payoff differences in (1) can also be written as $u - u^k \in \mathcal{U}$.) We can now state formally our definition of cost-benefit representation:

Definition 2 (Cost-Benefit Representation)**.** *The agent's reasoning admits a* cost-benefit representation *if, for any $C \in \mathcal{C}$, there exist functions (c, V) and (W^0, W^1) s.t.:*

- **(Q.0)** $W^0 : \mathcal{U} \times \mathbb{N} \to \mathbb{R}$ *is such that, for any* $k \in \mathbb{N}$, $u \succsim_k^0 v$ *if and only if* $W^0(u,k) \geq W^0(v,k)$, *and such that* $W^0(u,k) = W^0(u^k, k)$;

- **(Q.1)** $W^1 : \mathcal{U} \times \mathbb{N} \to \mathbb{R}$ *is such that, for any* $k \in \mathbb{N}$, $u \succsim_k^1 v$ *if and only if* $W^1(u,k) \geq W^1(v,k)$, *and such that* $W^1(u,k) \geq W^1(u^k, k)$;

- **(Q.0-1)** $V : \mathcal{U} \times \mathbb{N}_+ \to \mathbb{R}_+$ *and* $c : \mathbb{N}_+ \to \mathbb{R}_+ \cup \{\infty\}$ *are such that, for any* $k \in \mathbb{N}_+$, *for any* $(u,E) \in C$, *and* $t \in \mathbb{R}$: $(u+t,1) \succsim_{k-1} (u,0)$ *if and only if* $V(u,k) + t \geq c(k)$, *and: (i)* $V(u,k) = W^1(u - u^{k-1}, k-1)$, *(ii)* $V(u,k) \geq 0$ *for all* $(u,E) \in C$, *(iii)* $V(u,k) = 0$ *if* u *is constant in* a, *and (iv) for any* $k \in \mathbb{N}_+$, $V(u,k) \geq V(u',k)$ *if* $\left(u(a,\omega) - u(a^{k-1},\omega)\right)_{(a,\omega)\in A\times\Omega} \geq \left(u'(a,\omega) - u'(a^{k-1},\omega)\right)_{(a,\omega)\in A\times\Omega}$.

We say that a (c,V)*-representation is* unique *if, for any cognitive equivalence class* $C \in \mathcal{C}$, *whenever* (V,c) *and* (V',c') *represent the agent's depth of reasoning, then* $c = c'$ *and* $V(\cdot,k) = V'(\cdot,k)$ *for all* k *such that* $c(k) \neq \infty$.

In words, for any choice problem (u,E), and for any step k, there are functions W^0 and W^1 which represent the agent's assessment of his *current* and *future understanding*, and the agent keeps thinking as long as the value of reasoning $V(u,k)$ at step k is higher than the cost $c(k)$. If the agent receives an extra t to think about u, then our representation linearly separates $V(u,k)$ from t, so that he continues reasoning if $V(u,k) + t > c(k)$. Moreover, the value of reasoning satisfies the following properties: (i) it is equal to the perceived gain in payoff from switching from the current action a^{k-1} to what the agent expects to learn, based on his assessment of his future understanding; (ii) it is never negative; (iii) it attains its minimum for payoff functions which are constant in a (for which there is no instrumental value of reasoning); and (iv) it is (weakly) increasing in the payoff differences, which provide a minimal notion of the potential payoff gains. These properties are minimal for an *instrumental* notion of value of reasoning, and in this sense this definition provides a minimal notion of cost-benefit representation.

The restrictions on the W^0 and W^1 functions in points (Q.0) and (Q.1) also express ideas consistent with instrumental reasoning. Condition (Q.0) states that the agent's assessment of his current understanding only depends on the payoff consequences of the current action, a^k. Condition (Q.1) states that the agent's assessment of his future understanding is such that he never strictly prefers to committing to the current action a^k and ignoring what he might learn.

The separation between V and c is what makes the representation meaningful. The role of the cognitive equivalence classes (Def. 1) is that they allow for simple comparative statics. Specifically, the representation maintains that the cost function is the same for any choice problems (u,E) and (u',E') within a c.e. class. Hence, cognitively equivalent problems may only differ in their value of reasoning, i.e. in the incentives to reason they provide. This representation therefore enables simple comparative statics within each cognitive equivalence class, as only the value of reasoning is shifted as payoffs vary. The manner in which this may occur will be discussed in Section 3.

Finally, note that the notion of uniqueness provided in Definition 2 is the strongest possible for the (c,V)-representation: if the cost of reasoning is infinite at some step, then the agent would choose not to reason for any specification of the value of reasoning.

3 Properties of the Reasoning Process and Representations

In this section we introduce the properties of the reasoning process – formally, axioms on the preference relation $\succsim_s$ – and provide various representation theorems. We first provide the formal condition on cognitive equivalence classes we mentioned previously. It states that, holding E fixed, multiplying all payoffs by a positive scalar, or adding a constant to them, does not change the way in which the agent reasons about the problem.

Condition 1. *For any E, if u and v are such that, for some $\alpha \in \mathbb{R}_+$ and $\beta : \omega \to \mathbb{R}$, $v(a,\omega) = \alpha u(a,\omega) + \beta(\omega)$ for all (a,ω), then $(v,E) \in C(u,E)$.*

This Condition will be maintained throughout. This is a natural assumption for our consequentialist approach, and for some of the properties of reasoning we consider (namely, Properties 8 and 9 in the next section), which maintain a von Neumann-Morgenstern flavor. This condition plays no role in our Core Representation Theorem (which does not invoke Properties 8 and 9), but it is restrictive, and it may be desirable to modify it in future work. One can imagine, for instance, that if incentives were to increase dramatically, then the agent might approach the problem entirely differently. In that case, Condition 1 would have to be modified, for instance assuming that it only holds for multiplications by a scalar below some threshold $\bar{\alpha} > 0$. With that change, a similar modification of Property 8 would deliver for instance a result analogous to Theorem 2. This variation can be easily cast within our framework, but we ignore it for simplicity.

3.1 Core Properties and Representations

The first three properties are standard monotonicity and continuity properties:

Property 1 (Monotonicity)**.** *For each $s \in S$, and $x = 0,1$: If $u \geq v$, then $u \succsim_s^x v$. If $u >> v$ then $u \succ_s^x v$.*

Property 2 (Archimedean)**.** *For each $s \in S$, $x = 0,1$ and $v,u,w \in \mathcal{U}$ such that $u \succ_s^x v \succ_s^x w$, there exist $0 \leq \beta \leq \alpha \leq 1$ such that $\alpha u + (1-\alpha) w \succ_s^x v \succ_s^x \beta u + (1-\beta) w$.*

Property 3 (Continuity)**.** *For each $s \in S$ and $u \in \mathcal{U}$, if there exists $t \in \mathbb{R}$ s.t. $(u+t,1) \succsim_s (u,0)$, then there exists $t^* \leq t$ such that $(u+t^*,1) \sim_s (u,0)$.*

Since payoffs are already scaled in utils, the first two properties are natural in this context. The third property is also particularly weak. For instance, it allows lexicographic preferences such as $(u,0) \succ_s (u+t,1)$ for all u and t, which means that the agent cannot be incentivized to perform the next step of reasoning, no matter how high t is.

The next property defines the scope of our theory. Part (S.1) states that the reasoning process is purely instrumental in informing the player's choice. Thus, if i's payoffs are constant in a, the agent would never strictly prefer to think harder. Part (S.2) pushes the idea further, requiring that the incentives to reason are completely driven by the payoff differences between actions: if u and v are such that $u(a,\omega) - u(a',\omega) = v(a,\omega) - v(a',\omega)$ for all a, a' and ω, then u and v provide the agent with the same incentives to reason.

Property 4 (**Scope**)**.** *For each $s \in S$:*

S.1 If u is constant in a, then $(u, 0) \succsim_s (u, 1)$.

S.2 If $u - v$ is constant in a, then $(u, 0) \succsim_s (u, 1)$ if and only if $(v, 0) \succsim_s (v, 1)$.

An agent who is observed strictly preferring to think further for a constant u is in violation of S.1 (see Section 5.2). S.2 instead is violated if the agent thinks further in one problem than in another one whose payoffs only differ for a constant $t \cdot \mathbf{1}$.

The next axiom is a 'cost-independence' condition, and will be used as a calibration axiom to pin down the cost of reasoning:

Property 5 (Cost-Independence)**.** *For each $s \in S$, for any u, v that are constant in a, and for any $t \in \mathbb{R}$, $(u + t, 1) \sim_s (u, 0)$ if and only if $(v + t, 1) \sim_s (v, 0)$.*

To understand this property, first note that (S.1) implies $(u, 0) \succsim_s (u, 1)$ whenever u is constant in a, since thinking has no instrumental value. Suppose the agent can be made indifferent between thinking and not, if the former were accompanied by an extra reward $t \geq 0$ (that is, $(u + t, 1) \sim_s (u, 0)$). Then, that reward would also make him indifferent between reasoning or not given any other constant function. This condition is weak because it is only required for constant payoff functions. Also note that there may not be a reward t that can induce the agent to think, as would be the case for an absolute cognitive bound at mental state s.

Recall that we use a^s to refer to the action at an arbitrary mental state s, and a^k for the action at the k-th state s^k along a specific path. For each $u \in \mathcal{U}$ and s, we define u^s as $u^s(a, \omega) = u(a^s, \omega)$ for all (a, ω), and similarly write u^k instead of u^{s^k}. For instance, in the leading example with matrix u replicated below, if $a^{s^k} = a^k = a_1$, then

$$u = \begin{array}{|c|c|c|c|} \hline & \omega_1 & \omega_2 & \omega_3 \\ \hline a_1 & 1 & 0 & 0 \\ \hline a_2 & 0 & 1 & 0 \\ \hline a_3 & 0 & 0 & 1 \\ \hline \end{array} \qquad u^{s^k} = u^k = \begin{array}{|c|c|c|c|} \hline & \omega_1 & \omega_2 & \omega_3 \\ \hline a_1 & 1 & 0 & 0 \\ \hline a_2 & 1 & 0 & 0 \\ \hline a_3 & 1 & 0 & 0 \\ \hline \end{array} .$$

The next property states that, given that the agent would choose a^s if he stopped reasoning at s, at that mental state he is indifferent between the original payoff function u and u^s. Any agent who only cares about his final payoffs should satisfy this property.

Property 6 (Consequentialism)**.** *For any $s \in S$ and for any u: $u \sim_s^0 u^s$.*

The next property requires that, for any problem and for any mental state associated with it, the agent does not strictly prefer to commit to ignore what he might learn:

Property 7 (No Improving Obstinacy)**.** *For any (u, E) and $s \in S(u, E)$: $u \succsim_s^1 u^s$.*

As shown by Example 1 in Appendix A, this is an important property for a cost-benefit representation in that it guarantees that the value of reasoning is non-negative. If Property 7 is violated then at some mental state s the agent feels that, were he to think further, he would be better off committing to action a^s (equivalently: being rewarded by u^s) rather than following his future understanding. In this sense, violations of this property entail a preference for committing to disregard what might be learned.

Any standard model would trivially satisfy this property. Violations of this axiom entail an extreme lack of confidence in one's own reasoning, and may be associated with phenomena of 'thinking aversion'. These violations are discussed in Section 5.2.

Theorem 1 (Core Representation). *The agent's reasoning satisfies Properties 1-7 if and only if it admits a* unique *cost-benefit representation (cf. Def. 2).*

As discussed in Section 2.2, this result provides a minimal cost-benefit representation: the value of reasoning is never negative, and attains its minimum for payoff functions for which reasoning presents no instrumental value. Moreover, it is (weakly) increasing in the payoff differences, which provide a minimal notion of the potential payoff gains. We note that this representation also implies that $V(u,k)=0$ if a^{k-1} is optimal at all states.[3]

We also note that part (i) of the (Q.0-1) condition in Definition 2 also implies that it is possible to rank payoff functions by the incentives to reason they provide, using the $\succsim_k^1$ relation alone: for any cognitive equivalence class, we say that u provides stronger incentives to reason than v at step k if $(v+t,1)\succsim_k (v,0)$ implies $(u+t,1)\succsim_k (u,0)$ for all $t\in\mathbb{R}$. Under the representation, this is possible if and only if $V(u,k+1)\geq V(v,k+1)$, which (by part (Q.0-1.i) in Def. 2) is equivalent to $W^1\left(u-u^k,k\right)\geq W^1\left(v-v^k,k\right)$. But since $W^1(\cdot,k)$ represents $\succsim_k^1$, we have:

Corollary 1. *Fix a cognitive equivalence class, and let k be s.t. $c(k+1)<\infty$. Then, u provides stronger incentives to reason than v at step k if and only if $\left(u-u^k\right)\succsim_k^1\left(v-v^k\right)$.*

As emphasized in Section 2.2, the cost-benefit representation in Theorem 1 is uniquely pinned down by each equivalence class. Comparative statics are therefore simple within each class, in that two cognitively equivalent problems (u,E) and (u',E') are associated with the same cost of reasoning, and may only differ in the values $V(u,k)$ and $V(u',k)$. The little structure on V, however, limits the applicability of Theorem 1: instances in which payoffs can be transformed to ensure that the payoff differences increase uniformly across a and ω are rare. In practice, an intuitive notion of varying the stakes is to multiply payoffs by a scalar $\alpha\in\mathbb{R}_+$: if $\alpha>1$ (resp., $\alpha<1$), it is natural to think of the transformation as increasing (resp., decreasing) the stakes. But if $(u(a,\omega)-u(a^s,\omega))_{(a,\omega)\in A\times\Omega}$ has both positive and negative entries, then $(\alpha u(a,\omega)-\alpha u(a^s,\omega))_{(a,\omega)\in A\times\Omega}$ is not uniformly above or below $(u(a,\omega)-u(a^s,\omega))_{(a,\omega)\in A\times\Omega}$. Hence, the representation in Theorem 1 does not record this payoff transformation as increasing the incentives to reason. The next property characterizes a representation which allows these easy-to-implement comparative statics.

Property 8 (Payoff Magnification). *For any $s\in S$, if $u\sim_s^1 t\cdot\mathbf{1}$ for some $t\geq 0$, then $\alpha u\succsim_s^1 t\cdot\mathbf{1}$ for all $\alpha\geq 1$.*

[3] This follows from the the fact that payoff differences are negative if a^{k-1} is dominant, and from the result that V is decreasing in the payoff differences and achieves its minimum if they are all zero.

This is a natural restriction: If the agent expects that applying the insights of future understanding to payoff function u is as valuable as receiving a certain payoff $t \geq 0$, then it would be at least as valuable if those payoffs were magnified by a constant (recall that payoffs are already in utils). As further discussed in Section 5, violations of this property are associated with phenomena of 'choking' (e.g., Ariely et al. (2009)) and to 'fear of failure'.

Theorem 2 (Core* Representation)**.** *Under the maintained assumptions of Theorem 1, Property 8 is satisfied if and only if V is such that, for any $(u, E) \in C$ and $\alpha \geq 1$, $V(\alpha u, k) \geq V(u, k)$.*

Since the conditions for this representation are weak, it reveals that increasing incentives to reason by magnifying payoffs, which is easily implemented in experimental settings, is often justified.[4] In closing, we note that while its full characterization is beyond the scope of this paper, the value function for the fully sophisticated, forward-looking Bayesian agent, is consistent with Theorem 2 (and hence with Theorem 1). This can be seen in that all the properties of the V function in this theorem (namely, that it is never negative, it is 0 if u is constant in a, and $V(\alpha u, k) \geq V(u, k)$ for $\alpha \geq 1$) are satisfied by the value function of such an agent.

3.2 Expected Value of Reasoning

In this section we investigate properties which lead to an 'expected value' representation of the value of reasoning. From a formal viewpoint, the first such property is a standard independence axiom for the binary relation $\succsim_s^1$:

Property 9 (1-Independence)**.** *For each $s \in S$: For all $u, v, w \in \mathcal{U}$, $u \succsim_s^1 v$ if and only if $\alpha u + (1-\alpha) w \succsim_s^1 \alpha v + (1-\alpha) w$ for all $\alpha \in (0,1)$.*

Joint with monotonicity (Property 1), this property implies Property 8. Hence, at a minimum, adding this property to those of Theorem 1 ensures a value of reasoning with the properties stated in Theorem 2. More broadly, independence as usual ensures that beliefs can be represented in probabilistic terms. There is, however, one important difference relative to standard applications of independence: The beliefs to be elicited here are not only over the states ω, but also over the action a that the agent expects to choose at the next step of reasoning.

Intuitively, one would expect a certain consistency between an agent's view on his future beliefs over ω, and his outlook over his future action. This consistency does not come from independence alone, but it will be the subject of the next property, which imposes a minimal notion of *aptness* for the problem at hand. In words, the next property requires that, if mental state s belongs to the reasoning process about (u, E), then the agent expects that his future understanding – whatever it might be – would be at least as apt to the actual problem than to a similar one, in which one of the alternatives has been replaced by one of the others. Formally: for any $a', a'' \in A$, let the payoff function $u_{a' \to a''}$ be identical to u except that $u_{a' \to a''}(a', \omega) = u(a'', \omega)$ for any ω.[5] Then:

[4]If Condition 1 were modified in the way discussed in p. 11, then the statement of Theorem 2 would be modified to having $\alpha \in [1, \bar{\alpha})$ rather than $\alpha \geq 1$, where $\bar{\alpha}$ is the threshold previously referred to.

[5]Formally: for any $u \in \mathcal{U}$ and $a, a' \in A$, $u_{a' \to a''} \in \mathcal{U}$ is such that $u_{a' \to a''}(a, \omega) = u(a, \omega)$ if $a \neq a'$, and $u_{a' \to a''}(a', \omega) = u(a'', \omega)$. Effectively, it is as if act a' is replaced by a copy of another act a'' in the menu.

Property 10 (No Improving Replacement). *For any (u, E) and $s \in S(u, E)$, $u \succsim_s^1 u_{a \to a'}$ for all $a, a' \in A$.*

This property is extremely weak and intuitive. To see this, consider the problem in the leading example of Section 2, and let a_3 'become' a copy of a_1:

u

	ω_1	ω_2	ω_3
a_1	1	0	0
a_2	0	1	0
a_3	0	0	1

$u_{a_3 \to a_1}$

	ω_1	ω_2	ω_3
a_1	1	0	0
a_2	0	1	0
a_3	1	0	0

If $u_{a_3 \to a_1} \succ_s^1 u$, so that Property 10 is violated, then the agent expects his future understanding to yield higher payoffs in $u_{a_3 \to a_1}$ than in u. But since $u_{a_3 \to a_1}$ is the same as u but with a_3 replaced by a_1, it is as if the agent expects to regret following his own reasoning, should it suggest a_3, because he would prefer a_1 instead. Property 10 ensures that no such regret is expected, for any action. Hence, Property 10 imposes a very weak notion of aptness, and does not directly require optimality of the reasoning process. Yet, together with the other properties, it yields a representation in which, at each stage of reasoning k, it is *as if* the agent expects that if he were to further pursue his reasoning, with some probability $p^k(\mu)$ he would form new beliefs $\mu \in \Delta(\Omega)$, and that in each case he would choose an optimal response $a^*(\mu)$ to those beliefs. Formally, for any $u \in \mathcal{U}$ and $\mu \in \Delta(\Omega)$, we let $BR^u(\mu) := \arg\max_{a \in A} \sum \mu(\omega) u(a, \omega)$. Then:

Theorem 3 (EV-representation). *Under the maintained assumptions of Theorem 1, Properties 9 and 10 are satisfied if and only if V in Theorem 1 is such that*

$$V(u, k) = \sum_{\mu \in \Delta(\Omega)} p^k(\mu) \sum_{\omega} \mu(\omega) \left[u(a^*(\mu), \omega) - u\left(a^{k-1}, \omega\right) \right], \tag{2}$$

where $p^k \in \Delta(\Delta(\Omega))$, $\mu \in \Delta(\Omega)$, and $a^*(\mu) \in BR^u(\mu)$ for all $(u, E) \in C$.[6]

3.3 Value of Information

The functional form of V in Theorem 3 is reminiscent of the standard notion of 'expected value of information'. But the beliefs in the representation are obtained from the relation $\succsim_{k-1}^1$, which need not relate to the determinants of the current action a^{k-1}. In contrast, standard models of information require that agents are Bayesian in the stronger sense that they use a single prior over everything that affects choices. As such, for a value of information representation, the beliefs that describe the outlook on future steps of reasoning should be consistent with current choice a^{k-1}. Formally, let the mean beliefs $\hat{\mu}^k \in \Delta(\Omega)$ be defined from the representation in Theorem 3 so that

[6] We note that Property 7 is not needed for this result, as it is implied by the other properties jointly. Also, since the beliefs over the future understanding are pinned down by $\succsim_{s^k}^1$, which is constant within each cognitive equivalence class, these properties jointly entail a bound on the richness of the c.e. classes.

$\hat{\mu}^k(\omega) = \sum_{\mu \in \Delta(\Omega)} p^k(\mu)\,\mu(\omega)$ for each $\omega \in \Omega$. Then, in a standard Bayesian model, $a^{k-1} \in BR^u(\hat{\mu}^k)$ for each $(u, E) \in C$.

To obtain this kind of dynamic consistency, we add one other property of the reasoning process, which requires that, from the viewpoint of the *future* understanding relation $\succsim_s^1$, it is best to commit to the *current* action a^s than to any other $\hat{a} \in A$.

Property 11 (1-0 Consistency)**.** *For all (u, E), $s \in S(u, E)$ and $\hat{a} \in A$, $u^s \succsim_s^1 u^s_{a^s \to \hat{a}}$.*

Intuitively, consider again the leading example, and suppose that a_1 is the current action, i.e., $a^s = a_1$. For $u_{a^s \to a_3}$, a_1 has been replaced by a_3, and hence u, $u_{a^s \to a_3}$, u^s, and $u^s_{a^s \to a_3}$ are:

u

	ω_1	ω_2	ω_3
$a^s = a_1$	1	0	0
a_2	0	1	0
a_3	0	0	1

$u_{a^s \to a_3}$

	ω_1	ω_2	ω_3
$a^s = a_1$	0	0	1
a_2	0	1	0
a_3	0	0	1

u^s

	ω_1	ω_2	ω_3
a_1	1	0	0
a_2	1	0	0
a_3	1	0	0

$u^s_{a^s \to a_3}$

	ω_1	ω_2	ω_3
a_1	0	0	1
a_2	0	0	1
a_3	0	0	1

If it were the case that, from his current perspective of his future understanding, $u^s_{a^s \to a_3} \succ_s^1 u^s$, then it would effectively mean that the agent strictly prefers committing to another action a_3 rather than to his current action a_1. This is ruled out by Property 11.

Theorem 4 (Value-of-Information Representation)**.** *Under the maintained assumptions of Theorem 3, Property 11 is satisfied if and only if the representation in Theorem 3 is such that, for each $C \in \mathcal{C}$, for each $(u, E) \in C$ and for each k, $a^{k-1} \in BR^u\left(\hat{\mu}^k\right)$.*

The representation in this theorem says that it is *as if* the agent holds beliefs $\hat{\mu}^k$, to which the current action is a best response, and expects the outcome of further reasoning to yield optimal responses to posteriors obtained from Bayesian updating using signals generated by a Blackwell experiment. This is a completely standard value of information for a Bayesian agent who expects the next step of reasoning to be conclusive, or the subsequent steps of reasoning to be costless.[7]

Property 11 requires that the agent's attitude towards the next step of reasoning be the same as that entailed by his current understanding. While natural in standard information problems, this

[7]The model by Gabaix and Laibson (2005), Gabaix et al. (2006), which mainly focuses on the problem of directed cognition, also maintains a similar assumption of myopia. For a fully sophisticated forward-looking agent, who takes into account all possible ways in which his future reasoning may unfold, the value of reasoning would be a value function in the Bellman sense, and hence also incorporates the future costs which may or not be paid in different contingencies. A much richer dataset would be required to obtain such a representation, since one would need to disentangle the agent's beliefs about how his future reasoning unfolds over different steps (a note on this point is available from the authors upon request).

property is too narrow to accommodate general reasoning processes, since it rules out that the agent may not be fully aware of the determinants of the reasoning he has yet to perform. A disconnect between current behavior and future understanding is typical of several domains of research, in which the sheer complexity of certain problems makes it difficult for a researcher to anticipate what he might learn.

In conjunction with these factors, ethical constraints, prudence or lack of confidence about the current understanding often prevent the process of discovery from exhibiting the kind of dynamic consistency entailed by Property 11. In medical research, for instance, current practices are not abandoned even when researchers are very optimistic about new experimental techniques. As a simple example, suppose that a medical doctor can choose one of three treatments, $\{a_1, a_2, a_3\}$, whose effectiveness depends on the state of the world $\omega \in \{\omega_1, \omega_2, \omega_3\}$. Action a_1, the *status quo* treatment, yields a payoff of 0.4 in every state, while a_2 and a_3 are experimental treatments which perform better but only in certain states of the world: a_2 yields a payoff of 0.5 in states ω_2 and ω_3, but 0 otherwise; a_3 yields a payoff of 1 if the state is ω_3, and 0 otherwise. The doctor has not yet ruled out ω_1, but he is very optimistic that, upon further research, he will prove that ω_3 is the true state. He may still recommend the conventional treatment a_1 and at the same time assess the value of further research as close to 0.6. That is because he expects to confirm his belief that ω_3 is the true state, in which case he would switch from a_1 to a_3.[8] Hence, it is *as if* his current $a^s = a_1$ is dictated by one set of beliefs, but his value of reasoning is dictated by another. This is at odds with Property 11. Note that, with these beliefs, a Bayesian researcher would choose a_3 right away and have no incentive to research further.

In summary, for settings in which Property 11 is appealing, the Value of Information representation of Theorem 4 is justified. For settings in which it is not, such as situations in which unawareness, circumspection or sensitivity to confidence are relevant factors, alternative representations for the value of reasoning should be explored. These include Theorem 3, or the 'maximum gain' representation (Theorem 5), which we introduce next.

3.4 Circumspection in Deliberation

In this section we provide a representation which is particularly useful in applications, and which results from appending one property to those of Theorem 3. This new property expresses the idea that the agent is particularly deliberative or 'attentive' when making his choice. In words, the property requires that the agent regards further introspection as weakly more valuable (in the sense of the 'stronger incentives to reason' of Corollary 1) in the current problem (u, E) than it would be in any similar problem in which the payoffs of a single outcome $x = (a, \omega)$ were replaced by those of some other outcome $y = (a', \omega')$. As above, for any such x, y, let $u_{x\to y}$ be identical to u except that $u_{x\to y}(a, \omega) = u(a', \omega')$.[9]

[8] We argue that this is descriptively plausible, whether or not it is considered normatively compelling. For discussions on why rationality need not entail Bayesianism when there is insufficient understanding, see Gilboa, Postlewaite and Schmeidler (2012) and Gilboa, Samuelson and Schmeidler (2013).

[9] Formally: for any $u \in \mathcal{U}$ and $x, y \in A \times \Omega$, $u_{x\to y} \in \mathcal{U}$ is such that $u_{x\to y}(a, \omega) = u(a, \omega)$ whenever $(a, \omega) \neq x$, and $u_{x\to y}(x) = u(y)$. Effectively, it is as if outcome x is replaced by a copy of outcome y.

Property 12 (Circumspection)**.** *For any (u,E) and for any $s \in S(u,E)$, $u - u^s \succsim^1_s u_{x\to y} - u^s_{x\to y}$ for all $x, y \in A \times \Omega$.*

This property formalizes a weak notion of 'attentiveness' of the decision maker. Despite its weakness, the property has remarkable consequences for the representation:

Theorem 5 (Maximum Gain)**.** *Under the maintained assumptions of Theorem 1, Properties 9 and 12 are satisfied if and only if V in Theorem 1 is such that, for any $(u,E) \in C$,*

$$V(u,k) = \max_{\omega\in\Omega} u\left(a^*(\omega),\omega\right) - u(a^{k-1},\omega).$$

In words, it is *as if* the agent values the next step of reasoning by looking at the highest opportunity cost that playing according to the current understanding, a^{k-1}, may entail. This is a simple and intuitive heuristic, which has the advantage that it only depends on the payoffs, with no extra subjective parameters such as the beliefs in Theorems 3 and 4. This representation therefore is easy to use and limits the degrees of freedom of the model, which is especially desirable in applications (see Alaoui and Penta (2016a)).

The representation in Theorem 5 is a special case of the one in Theorem 3, in which the decision maker has the strongest incentives to reason. In this sense, the agent is fully optimistic about the value of pursuing his reasoning further (or, equivalently, he is fully pessimistic about his *current* understanding). However, this representation is inconsistent with that of Theorem 4: in Theorem 5, it is *as if* the agent is certain that the result of the next step of reasoning will be to learn that the state is $\omega^* \in \arg\max_{\omega\in\Omega}[u\left(a^*(\omega),\omega\right) - u\left(a^{k-1},\omega\right)]$. But with these beliefs, the representation in Theorem 4 implies that a^{k-1} is itself a best response to ω^*, hence $u\left(a^*(\omega^*),\omega^*\right) = u\left(a^{k-1},\omega^*\right)$, which means that the value of reasoning is 0. Given the definition of ω^*, however, this is only possible if a^{k-1} is a best response to every $\omega \in \Omega$. Thus, Properties 11 and 12 are inconsistent with each other in non-trivial problems, in which the optimal action varies with the state.[10]

A useful feature of our framework is that it is amenable to variations, which may be desirable in specific settings. Alaoui and Penta (2018) explore several alternative representations, which include various forms of cost-benefit criteria in R&D applications, or endogenous level-k reasoning in games (Alaoui and Penta's (2016a), also discussed in the next section). These representations can be obtained through simple modifications of the axioms. In this respect, the representation theorems in this section are best seen as templates for obtaining appropriate representations by suitably adjusting the baseline properties to match the desiderata of the specific problem at hand.

4 Applications

In this section we discuss a few ways in which the model can be taken to the data. Section 4.1 illustrates applications of the cost-benefit model to test its joint implications within specific settings, and how it can be connected to the literature on response time. In Section 4.2 instead we discuss

[10] Note that Property 10 is implied by the representation in Theorem 5, hence the gap between the representations in Theorems 4 and 5 is due to Properties 11 and 12 alone.

how our axioms can be tested in some settings, and, concretely, the kind of laboratory experiment that would be used. We also discuss how the model could be used for estimation with field data, and the requirements on the datasets.

4.1 Empirical Tests of the Cost-Benefit Model

The most immediate way to test the model's prediction is to look at the effect that changing the stakes of a problem – holding constant the cognitive equivalence class – has on the number of steps of reasoning undertaken by the agent. This is best done in settings in which individuals' reasoning processes are well-understood and in which the steps of reasoning are easily identified. One such example is provided by level-k reasoning in games. Section 4.1.1 discusses the experimental results of Alaoui and Penta (2016a), which tested the implications of the cost-benefit approach in the context of a model of *endogenous level-k reasoning*. Section 4.1.2 presents an extension to accommodate response time data, and discusses the experimental results of Alos-Ferrer and Buckenmeier (2019).

4.1.1 Strategic Thinking and Endogenous Level-k

As in standard models of level-k reasoning (e.g., Nagel (1995), Crawford and Iriberri (2007), Costa-Gomes and Crawford (2006)), the path of reasoning in the model of Alaoui and Penta (2016a, AP) is determined by iterating players' best-responses. It is thus assumed that games with the same (pure action) best-response functions are cognitively equivalent.[11] AP consider two versions of the model. In the first, 'detail free' version, the only restrictions on the cost and benefit functions are those entailed by the core representation of Theorem 1.[12]

AP's experiment studies subjects' initial responses in the *acyclical 11-20 game*. This is a two-player game in which players simultaneously report an integer between 11 and 20; a player's payoff is equal to his own report, plus an extra 20 if it is exactly one less than his opponent's, and an extra 10 if the two are identical. One of the main advantages of this game is that, for a wide range of level-0 specifications, each number from 12 to 20 is uniquely associated with a unique level-k action in this game, and hence shifts of the distribution of play can be directly mapped to shifts of the associated behavioral level-k.[13]

AP's experiment includes treatments designed to test both specific belief-effects and incentives to reason within their endogenous level-k model. Here we discuss the latter factor, incentives to reason, which is the focus of the present paper. To test whether behavior is consistent with the model's predictions as incentives to reason increase, AP repeat all the (belief) treatments conducted with

[11] To apply the model of reasoning to strategic settings, let $A \equiv A_i$ denote the set of the agent's actions in the game; the context E specifies a set of opponents N, with actions $A_{-i} = \times_{j \neq i} A_j \equiv \Omega$ and payoff functions $u_{-i} : A_i \times \Omega \to \mathbb{R}^{|N|}$ and $u_i : A_i \times \Omega \to \mathbb{R}$. In standard models of level-k reasoning, agents go through a sequence of 'increasingly sophisticated' actions: for each i, and for given 'level-0 profile' $a^0 = \left(a_i^0, a_{-i}^0\right)$, i's path of reasoning $\{a_i^k\}_{k \in \mathbb{N}}$ is such that, for each k and j, a_j^k is the best response to a_{-j}^{k-1}. The path of reasoning is thus a sequence of mental states $(s_{(u,E)}^k)_{k \in \mathbb{N}}$ with associated actions $\{a_i^k\}_{k \in \mathbb{N}}$.

[12] More precisely, AP used a version of Theorem 1 in which $V(\cdot, k)$ is increasing in the positive payoff differences, which follows from a slight modification of the Scope Axiom (cf. Alaoui and Penta (2018)).

[13] The *acyclical 11-20 game* is a modified version of Arad and Rubinstein's (2012) 11-20 game. The difference is that Arad and Rubinstein's (2012) version does not include the extra reward in case of tie, and hence it does not allow a unique identification of subjects' levels from their reports.

the extra reward of 20 in the acyclical 11-20 game, with a 'high payoff' version in which the extra reward is increased to 80. This variation entails a change of the payoff function, from u to v, which increases the payoff differences and hence falls within the domain of the representation in Theorem 1. The experimental results show that, as payoffs are increased, the empirical distribution of action shifts towards lower numbers, thereby providing clear evidence that subjects' depth of reasoning increases as the incentives to reason are increased in the sense of the core representation Theorem 1 above (cf. footnote 12). The significance of these shifts are confirmed by OLS regressions, Wilcoxon signed-rank tests, and Kolmogorov-Smirnov equality of distribution tests, and are also consistent with individual level data (see Alaoui and Penta (2016a), online Appendix).[14]

The second version of the model applies the maximum gain representation of Theorem 5 to obtain sharper predictions. With this added structure, AP apply the model to the experiments in Goeree and Holt (2001, GH), and show by means of a calibration exercise that it delivers quantitatively accurate predictions across different games. In particular, exploiting the parsimony of the maximum gain representation, AP use a version of their endogenous level-k model with with a single free parameter, they calibrate it to match the data in *one* of GH's games, and find that the predictions of the calibrated model for *all* the other static games in GH are highly consistent with their experimental findings.

4.1.2 Endogenous Response Time

In many settings individuals' reasoning processes are not as well understood, or the mapping from choice to the step reached may not be as clear. A growing literature has thus focused on auxiliary measures of individuals' reasoning, such as *response time* and *attention allocation*. It seems thus tempting to identify changes in the depth of reasoning in our model with changes of response-time (or of the attention allocated) of an equal direction. But such a connection is not straightforward, since the model above does not refer directly to time. In particular, the unit of measure of a step of reasoning in our model is not time, but a 'unit of understanding', and the amount of time needed to achieve it may vary from case to case: if it is harder to reason about problem (u, E) than (u', E'), then any given 'unit of understanding' may take longer in (u, E), and hence (u', E') may induce a lower response time even if it were associated with a larger depth of reasoning. Based on this logic – and abstracting from the possibility of endogenous changes in the individual's *focus* – one should expect that if two problems are cognitively equivalent, and as such equally difficult to think about, then every step of reasoning would take an equal amount of time in the two problems. Hence, if (u, E) and (u', E') are cognitively equivalent, and the former induces a higher depth of reasoning than the latter, then we expect a higher response time in (u, E) than (u', E').

The connection between response time and depth of reasoning is less straightforward when problems are compared across cognitive equivalence classes, for the following reason: if (u, E) and (u', E') induce the same value of reasoning, but (u, E) has a higher cost of reasoning than (u', E'), then even though the number of steps would be lower in (u, E), each of them would take longer than in (u', E'),

[14]Further support to the detail-free model is provided by Alaoui, Janezic and Penta (2020), who test more subtle implications of AP's detail-free model, and by Esteban-Casanelles and Gonçalves (2020), who perform an experiment with payoff treatments similar to AP's, except that their transformations are based on rescaling the payoffs in the sense of the Payoff Magnification Property 8.

and the overall effect is indeterminate. Hence, our model suggests special caution in drawing inferences on subjects' depth of reasoning from response time data, when non-cognitively equivalent problems are compared.

To formalize these ideas, for simplicity we maintain the maximum gain representation from Theorem 5 throughout this section, and for any problem (u, E) and cost of reasoning $c : N \to \mathbb{R}_+$, we let $\mathcal{K}(u, E; c)$ denote the associated depth of reasoning.[15] As already discussed, the maximum gain representation is particularly convenient in applications, because the value of reasoning at every step is completely pinned down by the sequence $\{a^k\}_{k\in\mathbb{N}}$. For later reference, we say that two decision problems are *path-equivalent* if they induce the same sequences $\{a^k\}_{k\in\mathbb{N}}$; they are *cost-equivalent* if they are associated with the same cost-of-reasoning. Cognitively equivalent problems are always both path- and cost-equivalent; under the maximum gain representation, the converse holds too, and it is also possible to obtain unambiguous comparative statics on the depth of reasoning between path-equivalent problems (even if they are not cognitively equivalent), whenever the associated costs of reasoning are ranked uniformly across steps.

For any problem (u, E), let $\tau : \mathbb{N}_+ \to \mathbb{R}$ denote the *time function*, where $\tau(k)$ represents the time it takes to perform the k-th step of reasoning. Then, the *total response time* in problem (u, E), given the associated cost and time functions (c, τ), is equal to:

$$T(u, E; c, \tau) := \sum_{k=1}^{\mathcal{K}(u,E;c)} \tau(k).^{16}$$

It seems reasonable to assume that if it is harder to attain a certain understanding in one problem than in another, then not only is the cost of reasoning (weakly) higher, but it also takes (weakly) more time to attain that understanding. We thus connect our baseline model of reasoning to response time by means of the following simple assumption:

Assumption 1 (Difficulty)**.** *Let (u, E) and (u', E') be path-equivalent, with associated costs of reasoning c and c', and time functions τ and τ', respectively. Then, for every $k \in \mathbb{N}$: $c(k) \geq c'(k)$ implies $\tau(k) \geq \tau'(k)$.*

Together with Theorem 1, this assumption implies the following:

Remark 1. *If (u, E) and (u', E') are cognitively equivalent, and (u', E') has larger payoff differences, then $T(u', E') \geq T(u, E)$. If the two are not cognitively equivalent, then the comparison between $T(u', E')$ and $T(u, E)$ is indeterminate, even if they are path-equivalent and $c(k) \geq c'(k)$.*

A recent paper by Alos-Ferrer and Buckenmeier (2019) investigates the relationship between response time and level-k reasoning. The paper provides a theoretical model based on the response-time extension above, applied to the endogenous level-k setting of Alaoui and Penta (2016a), and

[15]Formally: $\mathcal{K}(u, E; c) = \min\{k \in \mathbb{N}_+ : c(k) \leq V(u, k) \text{ and } c(k+1) > V(u, k+1)\}$.

[16]The $T(\cdot)$ function corresponds to the *chronometric function* in the neuroeconomics and psycho-metric literature. This specification abstracts from the possibility that changes in stakes may have a direct impact on individual's *focus* (cf. Alos-Ferrer and Buckenmeier (2019)). This possibility could be accommodated in the model by letting $\tau(\cdot)$ be directly affected by the payoff function, independent of the cognitive equivalence class. Such an extension, however, would limit the possibility of unambiguous comparative statics on response time to the subset of cognitively equivalent problems which also induce the same focus.

an experiment designed to test the model's predictions. The main assumptions in the theoretical model are the following: (i) reasoning follows the standard level-k path of reasoning; (ii) the depth of reasoning is jointly determined by costs and benefits of reasoning, where the latter are related to the payoff structure of the game; (iii) given the resulting depth of reasoning, k, the associated deliberation time is given by the sum of the k per-step-deliberation times; and (iv) the per-step deliberation time decreases as the benefits of that step increase.

Compared to our model above, the key innovation is assumption (iv), motivated by an established phenomenon in the neuroeconomics and psychometric literature, which is that individuals take longer to choose between more similar alternatives. We note that this assumption effectively entails a restriction *across* cognitive equivalence classes. Under assumptions (i)-(iv), Alos-Ferrer and Buckenmeier (2019) make the following predictions: (1) the depth of reasoning is non-decreasing in the incentives; (2) given the incentives, response time is non-decreasing in the depth of reasoning; and (3) holding constant the depth of reasoning, deliberation time is non-decreasing in the incentives. The impact of increasing the incentives on response time therefore is indeterminate, since it leads to more steps of reasoning, but it reduces the time associated with each of them.

Alos-Ferrer and Buckenmeier's (2019) experiment is based on variations of well-known games in the level-k literature, in which various features of the payoff function are varied, so as to affect the incentives to think. These games include a version of Nagel's (1995) Beauty Contest, and variations of the 11-20 game along the lines proposed by Goeree et al. (2017). The main findings are: (a) The effect of higher incentives on the depth of reasoning is consistent with Prediction 1 in all games; (b) In standard "level-k" games, such as the Beauty Contest and the 11-20 games, there is strong evidence of a link between depth of reasoning and response time; (c) The effect of higher incentives on deliberation times varies across variations of the 11-20 game. When one also controls for the effect on depth of reasoning, there is evidence in support of assumption (iv), namely that the per-step deliberation time decreases with the incentives.

Besides confirming central insights of our theory, Alos-Ferrer and Buckenmeier's (2019) results are especially promising in that they identify features of individuals' reasoning processes across cognitive equivalence classes, on which very little is understood.

4.1.3 Remarks

The Endogenous Response Time (ERT) model above can be further extended to derive testable predictions for a recent experiment by Avoyan and Schotter's (2019, AS) on attention allocation in games.[17] The predictions for AS's experiment are obtained directly from the ERT model, simply by adding a straightforward connection between response time and time allocation and one assumption on the cognitive equivalence classes which formalizes precisely one of AS's working hypotheses. As we discuss in Alaoui and Penta (2018), this extension of the model accommodates *all* of the experimental results in AS, across all games in that experiment.

[17] In AS's *attention allocation task*, subjects are presented a pair of two-player games in matrix form, (G^1, G^2), and they are asked the fraction $\alpha^l \in [0,1]$ of a total (unspecified) time X to allocate to game G^l ($\alpha^1 + \alpha^2 = 1$). In AS's experiments, such games consist of variations of archetypal two-person games such as battle of the sexes, pure coordination, constant-sum games, etc.

Overall, the experimental results of Alos-Ferrer and Buckenmaier's (2019) and Avoyan and Schotter (2019) provide support to the joint implications of the model of Section 4.1.2, which combines the assumptions on the individuals' reasoning processes contained in Theorem 1, with the mild assumptions on the connection with response time (the Difficulty Assumption 1) and on the cognitive equivalence classes (which serve as identification restrictions, by determining the domain of applicability of the comparative statics). We note that the same joint assumptions – and particularly those on the cost-benefit representation – are also consistent with the experimental findings discussed in Section 4.1.1. Hence, the results from Alaoui and Penta (2016a), their applications to Goeree and Holt's (2001) experiments, the experiments in Alos-Ferrer and Buckenmaier (2019), Avoyan and Schotter (2019) Alaoui, Janezic and Penta (2020) and Esteban-Casanelles and Gonçalves (2020) overall form a rich set of experimental evidence in support of our model.

4.2 Discussion: Elicitation, Testability and Estimation

The previous section focused on the *joint* testability of the model. Here we discuss whether axioms can be tested *individually*, as well as matters of elicitability of the preference relations.

4.2.1 Preference Elicitation and Testability of Individual Axioms

Our model presumes a rich set of preferences, which may be difficult to elicit in some settings. In those cases, our model serves primarily to understand the properties that constitute cost-benefit in reasoning. However, as shown by the opening example, there are economically relevant settings in which such elicitation is not particularly problematic.[18] The same observations apply to any situation in which the steps of reasoning can be observed and kept fixed. For instance, consider a subject who faces the same payoff function u as in the leading example. In a lab setting, the three choices and states could be whether the majority of balls contained in an urn are red, green, or yellow. The deliberation process involves pushing a button on the computer, which will provide the subject with some information on the composition of the urn. In this context, we can inform the subject that the discovery process is fixed exogenously, but he is given no clue on what kind of information it will produce nor how. At each stage, we can ask him which action he believes best at any stage. We can vary the payoffs conditional on stopping, as well as if he continues, without changing the information process. We can ask if he prefers continuing and be rewarded according to any payoff function v, or stopping and be rewarded according to u. We can thus fully observe his preferences and directly test each of the axioms.

In other cases, we may not be able to directly control the deliberation process. For instance, suppose we give subjects a counting task, where they have to choose whether there are more red, green or yellow dots on a screen. They get 1 if correct and 0 if not. Suppose, however, that we can still interrupt them during their thinking (while we cannot 'force' a subject not to think, we can provide him with only a couple seconds to answer questions, so that his thinking would be very limited). Then, without knowing precisely at which k they have stopped or what path they follow, we can still test whether our axioms hold at that step: We can ask the subject what his current

[18] We leave aside the standard issues in decision theory with testing technical axioms such as continuity.

choice is; whether he wishes to have more time to think about the problem, to elicit $(u,1) \succsim_k (u,0)$ (or vice-versa); and also whether, conditional on not being given more time (i.e., having committed the current action a^k), he prefers to be rewarded according to u or another v. We can therefore get at least a partial test of our axioms in a simple way. To elicit more complex preferences in an incentive-compatible manner, the following lottery method can be used.

Lottery Method Elicitation. Here we use a concrete experimental example of this method, which we formalize in the online appendix. Staying with the counting task example, we now ask the subject the following question, at the stage at which we have stopped him, giving him again a few seconds to answer: "If you were to think more, there is one chance out of one million that, after you've thought, instead of giving you 1 if correct and 0 otherwise, in case (a) we pay you 2 if correct and 1 otherwise, and in case (b) we pay you 1.5 if correct and 1.4 otherwise. If you were to think more, would you prefer case (a) or case (b)?" The agent here has no incentive to change the way he reasons – surely, the likelihood of his obtaining case (a) (some payoff matrix u') or (b) (some payoff matrix u'') is so low that he might as well follow his initial reasoning process, about his original problem (which has payoff matrix u). He also no incentive to lie. His choice between (a) and (b), therefore, would reveal his preferences $u' \succsim_k^1 u''$, or vice-versa. The $\succsim_k^1$ preferences can hence be elicited in practice, although it is of course delicate in some settings.[19]

Now suppose we ask subjects the following question, under the same protocol: "You can choose between two cases: In case (a), if you were to think more, there is one chance out of one million that, after you've thought, instead of giving you 1 if correct and 0 otherwise, we pay you 2 if correct and 1 otherwise; in case (b), there is one chance in one million that instead of thinking more and receiving 1 if correct and 0 otherwise, your current action can't be changed, and you'd receive 1.5 if correct and 1.4 otherwise." Here too, the likelihood of receiving either u' (in case a) or u'' (in case b) is so low that we wouldn't expect the agent to change his reasoning on that account. He also has no incentive to lie. But if he prefers (a) to (b) then $(u',1) \succsim_k (u'',0)$ (and vice-versa). Hence, the 0-1 comparisons can be elicited in this manner, although this may be difficult if one cannot check that the agent actually does think more. In those cases, further integration of these ideas with more direct proxies for reasoning (such as response time or measures of neural activity) may prove fruitful. These are interesting questions for future research.

4.2.2 Estimation in R&D

We discuss next how future work could use our framework for estimation. For instance, consider again the leading R&D example of Section 2, and suppose that we have data over a collection of research tasks, indexed by $i = 1, ..., n$, for each of which we observe: (i) the number K_i of weeks/periods spent researching on it; (ii) payoffs $u_i(a,\omega)_{(a,\omega)\in A\times\Omega}$ which describe the profitability of the alternatives in task i at the various states; and (iii) for each week, the action/product which was considered best. Observations for different i may come from the same R&D department over different tasks, or from

[19]While more complex, this method bears some similarities with experimental procedures such as the widely used Multiple Price List method (e.g., Kahneman, Knetsch and Thaler (1990), Holt and Laury (2002), Andersen et al. (2006), Azrieli et al., (2018)), which relies on making choices with the knowledge that only a subset of the choices will be used for actual payments.

different firms within an industry. Note that, if we assumed for instance the representation in Theorem 5, then the value $V_i(k)$ in weeks $k = 1, ..., K_i$ is completely pinned down by the data, and for each task i the decision to continue in those weeks only depends on the costs. For a tractable statistical model, suppose for instance that the costs for the i-th task is equal to $\tilde{c}_i(k) = c(k) + \epsilon_{i,k}$, where $\epsilon_{i,k}$ are, for instance, i.i.d. draws from a zero-mean normal with standard error σ_ϵ. Then, for each k, costs $(\tilde{c}_i(k))_{i=1,...,n}$ are i.i.d. draws from a normal $\mathcal{N}(c(k), \sigma_\epsilon)$, and a standard maximum-likelihood estimation could be performed to estimate the average cost $c(k)$ of performing the k-th week of research on these tasks. This basic model could be modified by assuming a bounded error distribution, or extended to allow for heteroskedasticity across different firms, or across different k's, as well as correlation (for instance, between observations across k for the same task, or across tasks within the same firm, etc). It may also be based on less parsimonious representations, such as the ones of Theorems 3 or 4, albeit imposing an extra burden on the data.

Obviously, the point of this discussion is only to illustrate in what sense the model could be extended to investigate this kind of empirical questions. A fully satisfactory answer necessarily requires more research. Nonetheless, our framework opens novel directions of research for the important problem of better understanding firms' R&D processes.

5 Core Violations and Psychological Phenomena

The previous section has focused on the applicability of our framework when the cost-benefit representation holds; this section focuses on its usefulness in understanding patterns of choice for which it may fail. These are *choking under pressure* (Section 5.1), *thinking aversion* and *rumination* (Section 5.2). These behavioral patterns are well-known in psychology research, and their relevance to economic settings has been increasingly acknowledged. Using our approach, we show that different psychological mechanisms which have been proposed to explain these phenomena map to distinct violations of the assumptions of our model. Our framework thus serves to disentangle these hypotheses and design tests to contrast their empirical implications.

5.1 Choking Under Pressure

Increasing financial incentives does not always improve performance in cognitive tasks, and it is sometimes even detrimental, as with "choking under pressure" (see, e.g., Camerer and Hogarth (1999) and Ariely, Gneezy, Loewenstein and Mazar (2009)).[20] The recent literature has shown a detrimental effect of increasing incentives in choice problems and in cognitive tasks (see, for instance, Gneezy and Rustichini (2000), or the 'packing quarters' and 'adding' tasks in Ariely et al.(2009)). In this type of problems, it is reasonable to maintain that a deeper understanding per se is not detrimental to performance, and hence a worse performance is associated with a lower depth of reasoning in our model.[21]

[20] The papers above contain detailed discussions of this topic and the proposed mechanisms in the literature. There is also a vast psychology literature which we do not survey; see also Baumeister and Showers (1986) and Beilock, Kulp, Holt and Carr (2004).

[21] This assumption does not necessarily apply to tasks that do not involve actual reasoning or introspection, such as tasks based on motor skills (e.g., Neiss (1988)) and choices in which there is a detrimental effect of reflection

Mechanism	Property Violated
Cost-based explanations:	Condition 1 (c.e., cost)
Level effect (Anxiety)	β
Differences effect (Fear of failure)	α
Value-based explanations:	
Level effect (Distraction)	Scope Axiom 4-(S.2)
Differences effect (Self-handicapping)	Property 8 (Payoff magnification)
Path of Reasoning:	
Deflection	Condition 1 (c.e., path)

Table 1: *Choking under pressure.* From psychological mechanisms to violations.

A lack of impact of financial incentives is not at odds with cost-benefit per se, since the costs for the task at hand may be steep, and possibly infinite. But a detrimental effect is not consistent with our representation. Different mechanisms have been offered to explain performance declining with stakes (cf. Camerer and Hogarth (1999) and Ariely et al. (2009)). Using the key elements of our model, we classify these competing mechanisms into three categories: *Cost-based*, *value-based* and *(deflection of the) path of reasoning*. The precise mapping between each mechanism and the assumption of the model that it violates is discussed below, and summarized in Table 1. We will then discuss how an experiment can be designed to separate these different mechanisms.

5.1.1 Cost-based Explanations: Anxiety and Fear of Failure

One view is that the increased pressure of higher rewards takes up working memory, which makes reasoning more difficult. Within our model, an increase in reasoning difficulty translates to higher costs. Hence, even though the task to be completed remains exactly the same, the cost of reasoning increases as rewards increase. This implies that the reward itself changes the cognitive equivalence class, which is in violation of Condition 1.

But, by this condition, two choice problems (u, E) and (v, E) are cognitively equivalent if they differ only by an additive constant $(v = u(a, \omega) + \beta(\omega))$, or only by a multiplicative factor $(v(a, \omega) = \alpha u(a, \omega))$. This suggests that the general mechanism of increased pressure can be separated into two factors. A higher cognitive load can be due to the pure level effect of higher rewards, which would lead to a reduced performance even when payoffs are increased by a constant. Alternatively, it can be due to larger payoff differences between succeeding and failing. These explanations can therefore be distinguished with the appropriate experimental design. The level effect mechanism would lead to a reduced performance when payoffs $u(a, \omega) + \beta(\omega)$ are increased by a sufficiently large $\beta(\omega)$, while the differences effect mechanism would lead to reduced performance when payoffs $\alpha u(a, \omega)$ are magnified by a sufficiently large α. Whereas the terms are sometimes used with different meanings in the psychology literature, it is natural to associate the level effect mechanism to phenomena of *anxiety*, and the differences effect mechanism to *fear of failure*.

interrupting automatic processes - also referred to as the "centipede effect" (see Camerer, Lowestein and Prelec (2005) and references therein).

5.1.2 Value-based Explanations: Distractions and Self-handicapping

Another view is that higher incentives distract the individual from the task at hand, thereby diminishing his motivation. In our model, a reduction in motivation translates to a decrease in the value of reasoning. Paralleling the discussion on costs, this value-based explanation can also be decomposed into differences and pure level effects.

The first mechanism, in which the value is lower when the incentives should be higher, is a failure of Property 8: under this mechanism, reasoning about αu when $\alpha > 1$ has less value than reasoning about u, even though there is more to gain. This kind of violations may be non-monotonic in the increase in α, which would be consistent with the empirical findings. Gneezy and Rustichini (2000), for instance, find that performance first worsens from awarding zero to low monetary rewards, but then improves as rewards are higher. In contrast, in the Yerkes-Dodson law (Ariely et al., (2009)) the violation of Property 8 is in the opposite direction: performance first improves and then diminishes.

The second mechanism, in which adding a constant payoff reduces the value of reasoning (V is higher for u than for $u(a, \omega) + \beta(\omega)$, $\beta(\omega) > 0$), is a failure of Property 4 (S.2). Psychological mechanisms such as *self-handicapping* and *distraction* in this context map naturally to the differences and pure level effects, respectively.

Experimental design to identify violations. It is conceptually straightforward to disentangle level effects from difference. Distinguishing between cost-based and value-based explanations, however, may seem more challenging. In the absence of our framework, the parallel between the cost and value mechanisms creates difficulty in disentangling them, even conceptually. Our model can be used to achieve this objective, because these different components (Condition 1 (parts α and β), Property 4 and Property 8) are clearly distinct. We now discuss an experimental design to identify which of the cost- or value-based explanations would explain the observed departure from behavior consistent with the cost-benefit approach.

First, since the different explanations map to different axiom violations, these can be directly tested, for instance using the elicitation methods method discussed in Sections 2 and 4.2. But those techniques require a rich dataset; here we provide a simpler alternative. Consider a cognitive task, similar to the experiment of Dean and and Neligh (2017), in which a subject is asked whether a screen with n balls has more red or green balls (states $\omega = r, g$, respectively). Let u be a payoff matrix which pays 1 if the decision is correct, 0 otherwise. Now suppose that we increase all the payoffs to $u' = \alpha u$ or to $u'' = u + \beta$, where $\alpha > 1$ and $\beta > 0$, while leaving the task (and hence the environment E) constant:

$$u = \begin{array}{|c|c|c|} \hline & r & g \\ \hline a_1 & 1 & 0 \\ \hline a_2 & 0 & 1 \\ \hline \end{array} \quad u' = \begin{array}{|c|c|c|} \hline & r & g \\ \hline a_1 & \alpha & 0 \\ \hline a_2 & 0 & \alpha \\ \hline \end{array} \quad u'' = \begin{array}{|c|c|c|} \hline & r & g \\ \hline a_1 & 1+\beta & \beta \\ \hline a_2 & \beta & 1+\beta \\ \hline \end{array}$$

We maintain here that performance increases with depth of reasoning, as we would expect with such a task.[22] Since our aim is to disentangle different source of violations, we assume that if they

[22]For simplicity, suppose that there are no learning effects from task repetition across treatments, if the comparison is made within groups. Alternatively, these treatments can be compared across groups.

occur, they are due to either the cost-based or the value-based mechanisms, but not both. With this hypothesis, if performance increases going from u to u' but is the same in u and u'', then we cannot rule out that the individual is consistent with all the properties. But if performance decreases from u to u', then he is violating either Condition 1 (part-α), or Property 8. If performance changes from u to u'', then he is violating either Condition 1 (part-β) or Property 4-S.2 (see Table 1).

Consider first the case in which performance decreases from u to u'. To further separate whether it is Condition 1 (part-α) or Property 8 that is violated, the latter can then be tested through the elicitation method discussed in Section 4.2. If it holds, then it is Condition 1 that is violated. For the case in which performance decreases from u to u'', an additional treatment can be used to separate Condition 1 (part-β) from Property 4-S.2. The experiment consists of maintaining u fixed, but letting the environment vary, so that (u, E) can be compared to (u, E'), where environment E' is identical in every way to E except that the task is more complex. For instance, suppose that there are now $m > n$ balls to count, but otherwise maintaining the same composition of the urn. We make the identification restriction that this change increases the cost, and leaves all else identical. In that case, if the agent's performance increases (in addition to his performance having changed from u to u'' previously), then it must be that it is Property 4-S.2 that is violated.

5.1.3 Changing the path of reasoning: deflection

The discussions above focused on cost and value-based explanations. Here we focus on a distinct explanation, which is that "increased motivation tends to narrow individuals' focus of attention on a variety of dimensions [...] This can be detrimental for tasks that involve insight or creativity, since both require a kind of open-minded thinking that enables one to draw unusual connections between elements." (Ariely et al. (2009), p. 453). That is, the additional motivation affects how individuals think, which translates in our model to higher rewards changing the reasoning process itself. Formally, the reasoning process $(s^k)_{k\in\mathbb{N}}$ associated with the lower rewards is not the same as the reasoning process $(\hat{s}^k)_{k\in\mathbb{N}}$ associated with higher rewards. This is a separate violation of Condition 1, in that it is not necessarily the costs that change, but the path of reasoning itself.

Our framework is useful for distinguishing this mechanism from the previous ones. For instance, with the value-based explanations, the path of reasoning is *not* affected by the reward. In particular, consider the value-based mechanism, and suppose that gradually increasing payoffs has the non-monotonic effect of first decreasing and then increasing the value of reasoning, which would translate to the performance first worsening and then improving. But since, according to that mechanism, it is the value that changes and not the path, high stakes may make the individual revert to solutions from much lower rewards.[23] In contrast, if the reasoning process associated with higher stakes induces a different sequence of actions from the original one, as implied by the change in path mechanism, such a reversion need *not* be observed.

[23] As previously discussed, this non-monotonicity in performance would be consistent with Gneezy and Rustichini (2000). This reversion would still be observed for the range of intermediary rewards if, instead, incentives to reason first increase and then decrease, as is consistent with the Yerkes-Dodson Law.

Psychological Phenomenon	Property Violated
Thinking aversion	Property 7 (No Improving Obstinacy)
Rumination and worry	Scope Axiom 1-(S.1)

Table 2: From psychological phenomena to violation of properties

5.2 Thinking aversion and rumination

Other violations of cost-benefit analysis include thinking aversion and rumination (Table 2). What we call *thinking aversion* refers to the notion that the agent may find thinking to be detrimental. This is the case if, for instance, the agent fears he could not neglect a harmful outcome of reasoning. He might then want to commit not to reason, i.e., $u^s \succ_s^1 u$. But this is a direct violation of Property 7, and is therefore inconsistent with the representation theorems in which the value of reasoning is never negative.[24] *Rumination* and *worry* refer to reasoning about problems in a way that is not useful for choice, either because they are in the past (rumination) or because nothing can be done to alter the outcome (worry; see Nolen-Hoeksema, Wisco and Lyubomirsky (2008)). In other words, the agent keeps thinking even though reasoning is not instrumental. This means that even for some choice problem in which u is constant a, $(u, 1)$ could be strictly preferred to $(u, 0)$. This is a violation of Property 4-S.1, which can be tested directly by eliciting the agent's preference to think about payoff functions which are constant in a.

6 Conclusions

In this paper we have provided a foundation for a cost-benefit analysis in reasoning, and we have analyzed different properties of the reasoning process together with the representations that they entail. We have shown that the model delivers new testable predictions, we have illustrated that our framework can inform rigorous discussions of alternative criteria, and we have provided a few applications. We have also shown that our framework can be used to disentangle competing theories of psychological phenomena that are at odds with the cost-benefit approach. Here we discuss some directions of future research.

One direction is to explore further representations of the value of reasoning, following the route suggested in the paper. Another direction of research is more centered around *costs*. Understanding the determinants of the costs of reasoning, which are constant within each cognitive equivalence (c.e.) class, requires a deeper analysis of the cognitive partition. Definition 1 and Condition 1 provide basic restrictions for cognitive equivalence, but these conditions can be extended, along the lines suggested in Section 4.1.2. The extent to which these restrictions hold could also be investigated empirically, as can the shape of cost functions associated with each c.e. class. A partial ranking of cost functions across c.e. classes, as well as the extent to which such a ranking is robust across individuals, could also be identified. This ranking would then be of use to compare difficulties of choice problems, and

[24]This concept is distinct from the notion of *thinking aversion* considered by Ortoleva (2013), in which a decision maker may prefer smaller menus so as to avoid the cognitive costs associated with larger menus.

to shed light on properties of the cost functions.[25]

Yet another direction of research involves further integration of the ideas in this paper with neuroeconomics. As we discussed, a direct elicitation of the primitives of our framework is possible in some settings, but may be more problematic in others. In those cases, our axioms mainly serve to understand the primitive properties of cost-benefit in reasoning, and provide a language to identify the cognitive building blocks of reasoning. More direct proxies for reasoning (such as response time or measures of neural activity), however, may help overcome some of these limitations, and provide more direct ways of testing our axioms. They may also serve to gain insight across c.e. classes, as in the case of Alos-Ferrer and Buckenmeier (2019) discussed in Section 4.2.2. Brain activity and response times can also be useful to elicit the degree of cognitive effort and deliberation, and our model can inform the design of experiments inspired by comparative statics on deliberation as incentives vary. In this vein, it would be interesting to explore the connection between our approach and neuroeconomics models of choice (e.g., Bogacz et al. (2006), Padoa-Schioppa and Rustichini (2014), and Rustichini and Padoa-Schioppa (2015), among others). Combining neuroeconomics with our model, and the axiomatic approach more broadly, also has the potential to identify c.e. classes and to analyze different dimensions of cognitive sophistication.

From a more applied perspective, ideas of bounded depth of reasoning have recently been applied to market settings as well. For instance, Hortacsu et al. (2016) apply Camerer et al.'s (2003) cognitive hierarchy model to firms' bidding behavior in electricity markets. They document that larger firms typically exhibit higher strategic sophistication, and they also find evidence that mergers endogenously increase firms' sophistication. Our model provides a framework to investigate the causes of both phenomena. For instance, an obvious explanation of the positive correlation between size and sophistication is that larger firms have lower 'costs of reasoning'. While this might be plausible, it is also possible that larger firms indeed may have higher costs. Our model suggests a more nuanced view. In particular, the larger firms' higher sophistication may be due precisely to the larger stakes of their choices which, given our representations, increase the value of reasoning. Different representations of the value of reasoning would thus provide different ways of connecting the underlying market structure to firms' strategic sophistication. This in turn can provide the extra structure to identify how much of the increase in strategic sophistication is due to internal characteristics of the firm (as in the 'lower cost' explanation), or to the change in the market structure, which affects the value of reasoning.

The research agenda discussed here is non-exhaustive, as the tradeoff between cognitive costs and value of reasoning is ubiquitous. There are various economic environments in which individuals reason in a stepwise fashion, and in which tradeoffs exist between cognitive costs and a value of reasoning. Our model provides a language for analyzing these diverse settings in a unified way.

[25]If cost functions c and c' associated with c.e. classes C and C', respectively, are such that $c(k) \geq c'(k)$ for all $k \in \mathbb{N}$, then we can say that choice problems in C are more difficult than problems in C'. See Mathevet (2014) for a classification of games by degree of complexity, based on an axiomatic approach.

Appendix

A Proofs

Proof of Theorem 1:

It will be convenient to introduce the following 'payoff differences' notation: for any $u \in \mathcal{U}$ and $\hat{a} \in A$, let $D(u, \hat{a}) : A \times \Omega \to \mathbb{R}$ be defined as

$$D(u, \hat{a})(a, \omega) = u(a, \omega) - u(\hat{a}, \omega) \text{ for all } (a, \omega) \in A \times \Omega. \tag{3}$$

We also let $D(u, k) := D(u, a^k)$. Fix the c.e. class $C \in \mathcal{C}$. For any k, the existence of continuous and monotonic functions $W^0(\cdot, k)$ and $W^1(\cdot, k) \in \mathbb{R}^{\mathcal{U}}$ which represent $\left(\succsim_k^0\right)$ and $\left(\succsim_k^x\right)$ follows immediately from the completeness, transitivity, continuity and monotonicity of the $\left(\succsim_k^x\right)_{x=0,1}$ relations.

Part 1 (Cost Identification):

Step 1.1: If $(u, 0) \succsim_k (u + t, 1)$ for all $u \in \mathcal{U}$ and $t \in \mathbb{R}$, then set $c(k+1) = \infty$. If not, let $\hat{u} \in \mathcal{U}$ be such that $(\hat{u} + t, 1) \succsim_k (\hat{u}, 0)$ for some $t' \in \mathbb{R}$. In that case, we show that

$$\exists c(k+1) \in \mathbb{R}_+ \text{ s.t. } u \in \mathcal{U}, \left(u^k + c(k+1), 1\right) \sim_k \left(u^k, 0\right) \text{ for all } u \in \mathcal{U}. \tag{4}$$

To this end, we first show that for the $\hat{u}$ above, $\exists c \in \mathbb{R}_+$ s.t. $\left(\hat{u}^k + c, 1\right) \succ_k \left(\hat{u}^k, 0\right)$: suppose not, then (using, in order, the hypothesis on $\hat{u}$, Property 6 and the absurd hypothesis), $(\hat{u} + t, 1) \succsim_k (\hat{u}, 0) \sim_k \left(\hat{u}^k, 0\right) \succsim_k \left(\hat{u}^k + c, 1\right)$ for all $c \in \mathbb{R}_+$, but $(\hat{u} + t, 1) \succsim_k \left(\hat{u}^k + c, 1\right)$ for all $c \in \mathbb{R}_+$ contradicts the monotonicity of $\succsim_k^1$ (Property 1). Now, since $\exists c \in \mathbb{R}_+$ s.t. $\left(\hat{u}^k + c, 1\right) \succ_k \left(\hat{u}^k, 0\right)$, Property 3 implies $\exists c^* \in \mathbb{R}_+$ such that $\left(\hat{u}^k + c^*, 1\right) \sim_k \left(\hat{u}^k, 0\right)$. We then set $c(k+1) \equiv c^*$. Note that, by Property 4.2, $\left(\hat{u}^k + c^*, 1\right) \sim_k \left(\hat{u}^k, 0\right)$ implies $\left(u^k + c^*, 1\right) \sim_k \left(u^k, 0\right)$ for all u.

The following characterization follows: for any $u \in \mathcal{U}$,

$$(u, 1) \succsim_k (u, 0) \text{ if and only if } (u, 1) \succsim_k \left(u^k + c(k+1), 1\right). \tag{5}$$

(The only if follows by transitivity, using the hypothesis, result (4) and Property 6.)

Step 1.2: Note that, for any $u \in \mathcal{U}$ and for any k, $D(u, k) = D\left(u - u^k, k\right)$. Hence, Property 4.2 implies

$$(u, 1) \succsim_k (u, 0) \text{ if and only if } \left(u - u^k, 1\right) \succsim_k \left(u - u^k, 0\right) \tag{6}$$

Using (5) and (6), it follows that, for any $u \in \mathcal{U}$,

$$(u, 1) \succsim_k (u, 0) \text{ if and only if } \left(u - u^k, 1\right) \succsim_k \left(\left(u^k - u^k\right)^k + c(k+1), 1\right),$$

where $\left(u - u^k\right)^k$ is defined such that, for every (a, ω), $\left(u - u^k\right)^k (a, \omega) = u\left(a^k, \omega\right) - u^k\left(a^k, \omega\right) \equiv 0$. Hence, we conclude that, for any $u \in \mathcal{U}$,

$$(u, 1) \succsim_k (u, 0) \text{ if and only if } \left(u - u^k\right) \succsim_k^1 c(k+1). \tag{7}$$

Part 2 (Value Identification): From part 1, we have two cases: (1) $(u, 0) \succsim_k (u + t, 1)$ for all $u \in \mathcal{U}$ and $t \in \mathbb{R}$, in which case we set $c(k+1) = \infty$. In this case, any $V : \mathcal{U} \to \mathbb{R}$ vacuously represents preferences $\succsim_k$: $(u + t, 0) \succsim_k (u, 1)$ if and only if $V(k+1) + t < c(k+1) = \infty$. (2) The complementary case is such that there exists $\hat{u} \in \mathcal{U}$ and t such that $(\hat{u} + t, 1) \succsim_k (\hat{u}, 0)$ and $c(k+1)$ such that $(u, 1) \succsim_k (u, 0)$ if and only if $\left(u - u^k\right) \succsim_k^1 c(k+1)$ for all $u \in \mathcal{U}$. We thus focus here on this second case.

Step 2.1: First we show that, for each $u \in \mathcal{U}$, there exists $t^u \in \mathbb{R}$ such that $(u + t^u, 1) \sim_k (u, 0)$. We consider different cases: (i) if $(u, 1) \sim_k (u, 0)$ then $t^u = 0$; (ii) if u is constant in a, then $t^u = c(k+1)$; (iii) If $(u, 1) \succ_k (u, 0)$, then by Property 3, there exists a t^u such that $(u + t^u, 1) \sim_k (u, 0)$, and by monotonicity (Property 1), $t^u < 0$; (iv) If $(u, 0) \succ_k (u, 1)$, then there exists sufficiently high $t^* \in \mathbb{R}_+$ such that, $(u + t^*, 1) \succsim_k (u, 0)$: suppose not, then by the absurd hypothesis, we have $(u, 0) \succ_k (u + t, 1)$ for all $t \in \mathbb{R}$, and by Property 6 we have $(u, 0) \sim (u^k, 0)$. But since u^k is constant, we know from the previous step that, for $t' > c(k+1) \geq 0$, $\left(u^k + t', 1\right) \succsim_k \left(u^k, 0\right)$. Hence, by the assumption that $\succsim$ admits a transitive and complete extension, $\left(u^k + t', 1\right) \succ_k (u + t, 1)$ for all $t \in \mathbb{R}$, which contradicts monotonicity (Property 1). Now, since $\exists t^* \in \mathbb{R}_+$ s.t. $(u + t^*, 1) \succ_k (u, 0)$, property 3 implies $\exists t^u \in \mathbb{R}_+$ such that $(u + t^u, 1) \sim_k (u, 0)$.

Step 2.2: Thus, $t^u \geq 0$ if and only if $(u, 0) \succsim_k (u, 1)$. Furthermore, by axiom 4.2, $t^u = t^{u-u^k}$, because $D(u, k) = D\left(u - u^k, k\right)$. Consider the following specification of the V function: for each $u \in \mathcal{U}$, let $V(u, k+1) = c(k+1) - t^u$. Now notice that $V(u, k+1) \geq V(v, k+1)$ if and only if $t^u \leq t^v$, but since $t^u = t^{u-u^k}$, $t^u \leq t^v$ if and only if $t^{u-u^k} \leq t^{v-v^k}$, that is if and only if $u - u^k \succsim_k^1 v - v^k$. Moreover, since $t^{u^k} = c(k+1)$ (because u^k is constant in a), we also have $V\left(u^k, k+1\right) = 0$. We show next that $V(u, k+1) \geq 0$ whenever $u \in \mathcal{U}(C)$: to see this, note that Property 7 implies that $u \succsim_s^1 u^k$ whenever $u \in \mathcal{U}(C)$, and since $u \in \mathcal{U}(C)$ implies $u - u^k \in \mathcal{U}(C)$ (by Condition 1), Property 7 also implies $u - u^k \succsim_k^1 u^k - u^k = \mathbf{0}$ (this is because $u^k - u^k = \left(u - u^k\right)^k$). Hence $V(u, k) \geq V\left(u^k, k\right) = 0$ for all $u \in \mathcal{U}(C)$.

Step 3 (Asymmetric Rewards): It remains to show that, for any t and u, $(u + t, 1) \succsim_k (u, 0)$ if and only if $V(u, k+1) + t \geq c(k+1)$. Given V defined above, clearly $V(u, k+1) + t \geq c(k+1)$ if and only if $t > t^u$. But monotonicity and the definition of t^u imply that $t > t^u$ if and only if $(u + t, 1) \succ_k (u + t^u, 1) \sim_k (u, 0)$.

Part 5 (Properties of W^0 and W^1): The fact that $W^0(u, k) = W^0(u^k, k)$ follows directly from Property 6. We show next that $W^1(\cdot, k)$ which represents $\succsim_k^1$ can be normalized to satisfy $W^1(u - u^k, k) = V(u, k+1)$. We proceed constructively: for any v such that $\exists u \in \mathcal{U} : v = u - u^k$, we let $W(v, k) = V(u, k+1)$ (note that if multiple $u, \hat{u}$ exists such that $v = u - u^k = \hat{u} - \hat{u}^k$, they induce the same value, because $V(u, k) := c(k) - t^u$ and $t^u = t^{u-u^k} = t^v = t^{\hat{u}-\hat{u}^k} = t^{\hat{u}}$). For any v for which $\exists u : v \sim^1 u - u^k$, we obviously set $W^1(v, k) = V(u, k+1)$. We thus need to extend the definition of W^1 to the other payoff functions v for which there is no u s.t. $v \sim_k^1 u - u^k$. There may me three cases:

1. $\exists u, \hat{u}$: $u - u^k \succ v \succ \hat{u} - \hat{u}^k$. Then, by the Archimedean property (Property 2), $\exists \alpha \in (0, 1)$ and $\tilde{v} := \alpha(u - u^k) + (1 - \alpha)(\hat{u} - \hat{u}^k)$ s.t. $v = \tilde{v}$. But notice that $\tilde{v}$ can be written as a difference, in particular $\tilde{v} = (\alpha u + (1 - \alpha)\hat{u}) - (\alpha u + (1 - \alpha)\hat{u})^k$, and hence we can set $W^1(v, k) = W^1(\tilde{v}, k)$.

2. $\forall u \in \mathcal{U}$, $v \in U^+ := \left\{v' \in \mathcal{U} : v' \succ u - u^k\right\}$. Let $w^{\sup} := \sup\left\{W^1\left(u - u^k\right) : u \in \mathcal{U}\right\}$. By continuity, there exists $\bar{u} \in \mathcal{U} : W^1\left(\bar{u} - \bar{u}^k\right) = w^{\sup}$. By continuity and monotonicity, for any $v \in U^+$, there exists a (unique) $\tau^v \in \mathbb{R}_+$ such that $v \sim_k^1 \left(\bar{u} - \bar{u}^k\right) + \tau^v$. Then, define $W^1(v) = w^{\sup} + \tau^v$. It is straightforward that for any $v, v' \in U^+$, $v \succsim_k^1 v'$ if and only if $\tau^v \geq \tau^{v'}$.

3. $\forall u \in \mathcal{U}$, $v \in U^- := \left\{v' \in \mathcal{U} : u - u^k \succ v'\right\}$. Let $w^{\inf} := \inf\left\{W^1\left(u - u^k\right) : u \in \mathcal{U}\right\}$. By continuity, there exists $\underline{u} \in \mathcal{U} : W^1\left(\underline{u} - \underline{u}^k\right) = w^{\inf}$. By continuity and monotonicity, for any $v \in U^-$, there exists a (unique) $\tau^v \in \mathbb{R}_+$ such that $v \sim_k^1 \left(\underline{u} - \underline{u}^k\right) - \tau^v$. Then, define $W^1(v) = w^{\inf} + \tau^v$. It is straightforward that for any $v, v' \in U^-$, $v \succsim_k^1 v'$ if and only if $\tau^v \leq \tau^{v'}$.

The W^1 thus defined represents $\succsim_k^1$, is continuous, and satisfies $W^1(u - u^k, k) = V(u, k+1)\, \forall u \in \mathcal{U}$.

Part 6 (Necessity of Properties 1-7): The existence of continuous functions $\left(W^0, W^1\right)$ representing $(\succsim_k^x)_{x=0,1}$ clearly implies that such binary relations are weak orders (maintained assumption) and satisfy the Archimedean Property 2. Property 1 also follows directly from the monotonicity of W^0 and W^1. Property 4 follows trivially from the representation; Property 3 is trivial; Property 5 follows in that for u and v constant, in the representation we have $t^u = t^v = c(k+1)$. Property 6 follows trivially from the stated property of W^0:

$W^0(u) = W^0(u^k)$. The necessity of Property 7 is due to the restriction $V(u, k+1) \geq 0$ for $u \in \mathcal{U}(C)$ in the theorem, as illustrated by Example 1 below.

Part 7 (Uniqueness) Suppose that (V', c') also represent the preferences. Then, for any $u \in \mathcal{U}$ and $t \in \mathbb{R}$,

$$V'(u, k+1) - c'(k+1) \geq t \iff V(u, k+1) - c(k+1) \geq t$$
$$\text{and } V'(u^k, k+1) = V(u^k, k+1) = 0$$

It follows that $c(k+1) = c'(k+1)$, and if $c(k+1) < \infty$, then $V'(u, k+1) = V(u, k+1)$ for every u. This completes the proof of Theorem 1. ■

Example 1 (Necessity of Property 7)**.** This example illustrates the necessity of Property 7 for a meaningful notion of value, which is non-negative and zero at constant payoff functions: Let $A = \{T, B\}$ and $\Omega = \{\omega\}$, so that we can omit ω in the following. Let the choice problem $(\hat{u}, E)$ be such that $\hat{u}(T) = 0$ and $\hat{u}(B) = -1$, and let the reasoning process be such that, for each k: (i) $a^k = T$; (ii) $\succsim_k^0$ and $\succsim_k^1$ are represented, respectively, by functions $W^0(u,k) = u(a^k)$ and $W^1(u,k) = \max_{a' \neq a^k} u(a')$; (iii) preferences $\succsim_k$ are such that $(u, 1) \succsim_k (u, 0)$ if and only if $V(u, k+1) \equiv W^1(u,k) - W^0(u,k) = V(u, k+1) = W^1(u - u^k, k) \geq c(k+1) > 0$. In this case, V does not reach its minimum at payoff functions that are constant in i's action, where it is equal to 0. To see this, let u be constant and identically equal to 0. Then, we have $V(\hat{u}, k) = -1 < V(u, k) = 0$. We now check that the pair (V, c) satisfies all the axioms and properties used in Theorem 1, except Property 7: (i) Property 4.1 is satisfied, since $V(u, k) = 0$ whenever u is constant, and hence $(u, 0) \succ_k (u, 1)$; 4.2 is satisfied, since $V(v, k) = V(u, k)$ if $v - u$ is constant in a; (ii) For any constant u, $(u + c(k+1), 1) \sim_k (u, 0)$, so Property 5 is satisfied; (iii) Function V is continuous, hence Property 3 is satisfied; (iv) Properties 1-6 are clearly satisfied. Property 7, however, fails: for $a^k = T$, $W^1(\hat{u}, k) = -1$ and $W^1(\hat{u}^k, k) = 0$, hence $\hat{u}^k \succ_k^1 \hat{u}$. □

Proof of Theorem 2:

Sufficiency: Relative to Theorem 1, we only need to show that, for any $u \in \mathcal{U}(C)$, $V(\alpha u, k) \geq V(u, k)$ whenever $\alpha \geq 1$. First note that Property 7 implies that $D(u, k) = u - u^k \succsim_k^1 u^k - u^k = \mathbf{0}$ (cf. Step 2.1 of the proof of Theorem 1). Then, by Properties 2 and 1, there exist $c^u \in \mathbb{R}_+$ such that $c^u \cdot \mathbf{1} \sim_k^1 D(u, k) \succsim_k^1 \mathbf{0}$ (the positiveness of c^u is due to monotonicity). Property 8 then implies that for $\alpha \geq 1$, $\alpha D(u, k) \succsim_k^1 D(u, k)$.

Now, let V be defined as in Step 2.1 of the proof of Theorem 1, then $V(\alpha u, k) \geq V(u, k)$ if and only if $t^u \leq t^{\alpha u}$ which is the case if and only if $t^{D(u,k)} \leq t^{D(\alpha u, k)} = t^{\alpha D(u,k)}$ (the latter equality follows trivially from the definition of D). This in turn is the case if and only if $\alpha D(u, k) \succsim_k^1 D(u, k)$, which is satisfied for $\alpha \geq 1$.

Necessity: By contradiction, suppose there exist $u \in \mathcal{U}(C)$, $c \in \mathbb{R}_+, \alpha \geq 1 : c \cdot \mathbf{1} \sim_k^1 u$ and $u \succ_k^1 \alpha u_i$. Now define $\hat{u}_i = u + u^k$, so that $D(\hat{u}, k) = u$, hence we have $D(\hat{u}, k) \succ_k^1 \alpha D(\hat{u}, k)$. By Theorem 1, $V(u, k) = W^1(u - u^k, k)$, that is $V(v, k) > V(u, k)$ if and only if $W^1(D(v, k)) > W^1(D(u, k))$. Hence, $D(\hat{u}, k) \succ_k^1 \alpha D(\hat{u}, k)$ implies $V(u, k) > V(\alpha u, k)$, contradicting the stated property of the value of reasoning.

Proof of Theorem 3:

Sufficiency: Given Theorem 1, we only need to show that for any k, the function $V(\cdot, k) : \mathcal{U} \to \mathbb{R}$ in that representation has the form in eq. 2. To this end, notice that with the addition of Property 9, the preference relation $\succsim_k^1$ satisfies the conditions for the mixture space theorem. Hence, for any $k \in \mathbb{N}$, there exists a function $\hat{W}^1(\cdot, k) : \mathcal{U} \to \mathbb{R}$ that represents $\succsim_k^1$ and satisfies $\hat{W}^1(\alpha u + (1-\alpha) v, k) = \alpha \hat{W}^1(u, k) + (1-\alpha) \hat{W}^1(v, k)$ for all $\alpha \in [0, 1]$ and $u, v \in \mathcal{U}$. Moreover, $\hat{W}^1(\cdot, k)$ is unique up to positive affine transformations. Because $\hat{W}^1(\cdot, k)$ is linear in $u \in \mathcal{U} = \mathbb{R}^{|A \times \Omega|}$, there exist $(\hat{\rho}(a, \omega))_{(a,\omega) \in A \times \Omega} \in \mathbb{R}^{|A \times \Omega|}$ s.t. $W^1(u, k) = \sum_{a,\omega} \hat{\rho}(a, \omega) \cdot u(a, \omega)$. By monotonicity, $\hat{\rho}(a, \omega) \geq 0$ for each (a, ω), and we define $\rho^k \in \Delta(A \times \Omega)$ normalizing such weights in the unit simplex, so that

$$\rho^k(a,\omega) = \begin{cases} \hat{\rho}(a,\omega) / \sum_{(a',\omega')} \hat{\rho}(a',\omega') & \text{if } \sum_{(a',\omega')} \hat{\rho}(a',\omega') > 0 \\ \frac{1}{|A\times\Omega|} & \text{otherwise} \end{cases}$$

Since this is a positive affine transformation, $W^1(u,k) = \sum_{a,\omega} \rho^k(a,\omega)\cdot u(a,\omega)$ also represents $\succsim_k^1$ by the uniqueness part of the mixture space theorem. For any such ρ^k, define $p^k \in \Delta(A)$ and $\mu = (\mu^a)_a \in \Delta(\Omega)^A$ as follows: for any $a \in A$, let $p^k(a) = \sum_\omega \rho^k(a,\omega)$ and define $\mu^a \in \Delta(\Omega)$ such that, for any ω,

$$\mu^a(\omega) = \begin{cases} \frac{\rho^k(a,\omega)}{p^k(a)} & \text{if } p^k(a) > 0 \\ \frac{1}{|\Omega|} & \text{otherwise} \end{cases}.$$

It is immediate to see that, for any (a,ω), $\rho^k(a,\omega) = p^k(a)\cdot\mu^a(\omega)$. Hence, without loss of generality we can represent $W^1(\cdot,k)$ as follows:

$$W^1(u,k) = \sum_a p^k(a) \sum_\omega \mu^a(\omega)\cdot u(a,\omega). \tag{8}$$

We show next that for each $u \in \mathcal{U}(C)$, $a \in BR^u(\mu^a)$ whenever there exists $\omega \in \Omega$ such that $\rho(a,\omega) > 0$. Suppose not. Then $\exists\hat{a}$ s.t. $\rho(\hat{a},\omega) > 0$ for some ω s.t. $\hat{a} \notin BR^u\left(\mu^{\hat{a}}\right)$. Then, let $a^* \in BR^u\left(\mu^{\hat{a}}\right)$ and consider payoff function $u_{\hat{a}\to a^*}$ (the payoff function identical to u except that the payoffs associated with action $\hat{a}$ are replaced by those of a^*). Then,

$$\begin{aligned} W^1(u_{\hat{a}\to a^*},k) &= \sum_{a,\omega} \rho^k(a,\omega)\cdot u_{\hat{a}\to a^*}(a,\omega) \\ &= \sum_{(a,\omega):a\neq\hat{a}} \rho^k(a,\omega)\cdot u(a,\omega) + \sum_{\omega\in\Omega} \rho^k(\hat{a},\omega)\cdot u(a^*,\omega) \\ &> \sum_{(a,\omega):a\neq\hat{a}} \rho^k(a,\omega)\cdot u(a,\omega) + \sum_{\omega\in\Omega} \rho^k(\hat{a},\omega)\cdot u(\hat{a},\omega) \\ &= W^1(u,k). \end{aligned}$$

But this conclusion contradicts Property 10. It follows that

$$W^1(u,k) = \sum_a p^k(a) \sum_\omega \mu^a(\omega)\cdot u(a^*(\mu),\omega) \text{ for all } u \in \mathcal{U}(C). \tag{9}$$

Notice next that functional $V(\cdot,k+1)$ in Theorem 1 represents the binary relation $\succsim^*$ defined as $u \succsim^* v$ if and only if $u - u^k \succsim_k^1 v - v^k$. Since $W^1(\cdot,k)$ represents $\succsim_k^1$, it follows that $u \succsim^* v$ if and only if $W^1\left(u - u^k, k\right) \geq W^1\left(v - v^k, k\right)$. Thanks to Property 9 we can thus set $V(u,k+1) = W^1\left(u - u^k, k\right)$. Since

$$W^1\left(u - u^k, k\right) = \sum_a p^k(a) \sum_\omega \mu^a(\omega)\cdot\left[u(a^*(\mu^a),\omega) - u\left(a^k,\omega\right)\right],$$

the representation follows from Theorem 1, noticing that $W^1\left(u^k - u^k, k\right) = 0$, and that (by Theorem 1) $u^1 \succsim_k u^0$ if and only if $u - u^k \succsim_k^1 c(k)$.

Necessity: Assume that

$$\begin{aligned} V(u,k+1) &= W^1\left(u - u^k, k\right) \\ &= \sum_a p^k(a) \sum_\omega \mu^a(\omega)\cdot\left[u(a^*(\mu^a),\omega) - u\left(a^k,\omega\right)\right]. \end{aligned}$$

Then, for any u, functions W^1 represents $\succsim_k^1$ and is such that

$$W^1(u) = \sum_{a \in A} p^k(a) \sum_{\omega \in \Omega} \mu^a(\omega) \cdot u(a^*(\mu), \omega).$$

It is immediate to verify that $\succsim_k^1$ satisfies Properties 9 and 10.■

Proof of Theorem 4:

Sufficiency: By contradiction, suppose there exists a' s.t. $\sum_\omega \hat{\mu}^k(\omega) u(a', \omega) > \sum_\omega \hat{\mu}^k(\omega) u(a^k, \omega)$. Then, substituting the definition of $\hat{\mu}^k$, we obtain

$$\sum_{\mu \in \Delta(\Omega)} p^k(\mu) \sum_\omega \mu(\omega) \left[u(a', \omega) - u(a^k, \omega)\right] > 0 \tag{10}$$

Define $v \in R(u, a^k)$ such that $v(a, \omega) = u(a, \omega)$ for all $a \neq a^k$ and $v(a^k, \omega) = u(a', \omega)$, we have $v^k \succ_k^1 u^k$ because eq. 8 represents $\succsim_k^1$. But $v^k \succ_k^1 u^k$ contradicts Property 11.

Necessity: it holds by construction, for the definition of $\hat{\mu}^k$.■

Proof of Theorem 5:

Sufficiency: From the proof of Theorem 3, it follows that under the assumptions of Theorem 1 plus Property 9, the value of reasoning takes the following form:

$$V(u, k+1) = W^1(u - u^k) = \sum_{(a,\omega)} \rho^k(a, \omega) \cdot \left[u(a, \omega) - u(a^k, \omega)\right].$$

We need to show that, adding Property 12, we have $\rho^k(a, \omega) = 0$ whenever $(a, \omega) \neq (a^*(\omega^*), \omega^*)$, where $\omega^* \in \arg\max_{\omega \in \Omega} u(a^*(\omega), \omega) - u(a^k, \omega)$. By contradiction, suppose that there exists $(a', \omega') \neq (a^*(\omega^*), \omega^*)$ s.t. $\rho^k(a, \omega) > 0$. Now, let v be such that

$$v(a, \omega) = u_{(a', \omega') \to (a^*(\omega^*), \omega^*)} = \begin{cases} u(a^*(\omega^*), \omega^*) & \text{if } (a, \omega) = (a', \omega') \\ u(a, \omega) & \text{otherwise} \end{cases}.$$

Then, by construction:

$$\begin{aligned} V(v, k+1) = W^1(v - v^k) &= \sum_{(a,\omega)} \rho^k(a, \omega) \cdot \left[v(a, \omega) - v(a^k, \omega)\right] \\ &= \rho^k(a', \omega') \left[u(a^*(\omega^*), \omega^*) - u(a^k, \omega^*)\right] + \sum_{(a,\omega) \neq (a'\omega')} \rho^k(a, \omega) \cdot \left[u(a, \omega) - u(a^k, \omega)\right] \\ &> \rho^k(a', \omega') \left[u(a', \omega') - u(a^k, \omega')\right] + \sum_{(a,\omega) \neq (a'\omega')} \rho^k(a, \omega) \cdot \left[u(a, \omega) - u(a^k, \omega)\right] \\ &= V(u, k+1) = W^1(u - u^k). \end{aligned}$$

Hence, $v - v^k \succsim_k^1 u - u^k$, which contradicts Property 12, since $v = u_{(a', \omega') \to (a^*(\omega^*), \omega^*)}$.

Necessity: trivial. ■

References

1. Alaoui, L. and A. Penta. 2016a. "Endogenous Depth of Reasoning," *Review of Economic Studies,* Vol 83, Issue 4 (2016): 1297-1333.

2. Alaoui, L. and A. Penta. 2016b. "Endogenous Depth of Reasoning and Response Time, with an application to the Attention-Allocation Task ," *mimeo*, UPF.

3. Alaoui, L. and A. Penta. 2018. "Cost-Benefit Analysis in Reasoning," Working Paper n. 1062 (version Oct. 2018), *Barcelona GSE.*

4. Alaoui, L. and A. Penta. 2021. "Cost-Benefit Analysis in Reasoning: The Value-Of-Information Case with Forward-Looking Agent," Working Paper n. 1785 (version July 2021), *UPF.*

5. Alaoui, L. K. A. Janezic and A. Penta. 2020. "Reasoning about Others' Reasoning," *Journal of Economic Theory*, Vol.189, 105091, August 2020.

6. Alos-Ferrer, C. and J. Buckenmeier. 2019. "Cognitive Sophistication and Deliberation Times," *mimeo.*

7. Aragones, E., I. Gilboa, A. Postlewaite and D. Schmeidler. 2005. "Fact-free Learning," *American Economic Review,* 95(5): 1355-1368.

8. Ariely, D., U. Gneezy, G. Loewenstein, N. Mazar. 2009. "Large Stakes and Big Mistakes," *Review of Economic Studies,* 76: 451-469.

9. Avoyan, Ala and Andrew Schotter (2019), "Attention in Games: An Experimental Study", forthcoming *European Economic Review.*

10. Bogacz R., Brown E., Moehlis J., Holmes P. and JD. Cohen. 2006. "The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks," *Psychological Review*, 113: 700-765.

11. Camerer, C. F. 2003. *Behavioral Game Theory,* Princeton Univ. Press, Princeton, NJ.

12. Camerer, C. F., G. Loewenstein and D. Prelec. 2005. "Neuroeconomics: How neuroscience can inform economics," *Journal of Economic Literature*, 34(1): 9-65.

13. Camerer, C. F. and R. Hogarth. 1999. "The effects of financial incentives in economics experiments: A review and capital-labor-production framework," *Journal of Risk and Uncertainty*, 19(1): 7-42

14. Caplin, A. and M. Dean. 2008. "Axiomatic Neuroeconomics," in: *Neuroeconomics: Decision Making and the Brain,* Chapter 3, P.W. Glimcher, C.F. Camerer, E. Fehr and R.A. Poldrack (Eds.), Academic Press, New York, 21–32.

15. Caplin, A. and M. Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review,* 105(7): 2183-2203

16. Costa-Gomes, M.A., and V. P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review,* 96(5): 1737-1768.

17. Crawford, V. P., M. A. Costa-Gomes, and Nagore Iriberri. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, 51(1): 5–62.

18. Crawford, V. P., and N. Iriberri. 2007. "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica,* 75(6): 1721-1770.

19. Dean, M. and N. Neligh. 2017. "Experimental Tests of Rational Inattention," *mimeo.*

20. De Finetti, B.. 1937. "La Prévision: Ses Lois Logiques, Ses Sources Subjectives," *Annales de l'Institut Henri Poincaré,* 7: 1-68.

21. Dekel, E., B. L. Lipman, and A. Rustichini. 1998. "Standard state-space models preclude unawareness," *Econometrica*, 66(1): 159–173.

22. Dekel, E., B. L. Lipman, and A. Rustichini. 2001. "Representing preferences with a unique subjective state space," *Econometrica,* 69(4): 891–934.

23. Dekel, E. and M. Siniscalchi. 2014. "Epistemic Game Theory" in: *Handbook of Game Theory*, vol. 4.

24. Esteban-Casanelles, T. and D. Gonçalves. 2020. "The Effect of Incentives on Choices and Beliefs in Game: An Experiment," *Mimeo.*

25. Gabaix, X. and D. Laibson. 2005. "Bounded rationality and directed cognition," *Mimeo.*

26. Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg. 2006. "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model,"*American Economic Review,* 96(4): 1043–1068.

27. Gigerenzer, Gerd and Wolfgang Gaissmaier. 2011. "Heuristic Decision Making," *Annual Review of Psychology,* 62:451–82.

28. Gilboa, I., A. Postlewaite, and D. Schmeidler.2012. "Rationality of Beliefs or: Why Savage's Axioms are neither Necessary nor Sufficient for Rationality," *Synthese,* 187, 11–31.

29. Gilboa, I., L. Samuelson and D. Schmeidler. 2013. "Dynamics of Inductive Inference in a Unified Framework, " *Journal of Economic Theory*, 148, 1399–1432.

30. Gneezy, U. and A. Rustichini. 2000. "Pay Enough or Don't Pay At All," *Quarterly Journal of Economics*, 115(3), 791–810.

31. Goeree, J. K., and C. A. Holt. 2001. "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review*, 91(5): 1402-1422.

32. Goeree, Jacob K., Philippos Louis and Jingjing Zhang. 2017. "Noisy Introspection in the 11-20 Game " *Economic Journal.*

33. Hortacsu, A., Luco, F., Puller, S. L. and D. Zhu. 2016. "Strategic Ability and Productive Efficiency in Electricity Markets," *mimeo*, Chicago.

34. Manzini, P. and M. Mariotti. 2007. "Sequentially Rationalizable Choice," *American Economic Review*, 97(5) 1824-1839.

35. Manzini, P. and M. Mariotti. 2014. "Stochastic Choice and Consideration Sets," *Econometrica*, 82(3): 1153-1176.

36. Mathevet, L. 2014. "Axiomatic Behavior in Repeated Games, *mimeo,* NYU.

37. Modica, S. and A. Rustichini. 1999. "Unawareness and Partitional Information Structures," *Games and Economic Behavior* 27: 265-298.

38. Nagel, R. 1995. "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85(5): 1313-1326.

39. Neiss, R. 1988. "Reconceptualizing Arousal: Psychological States in Motor Performance," *Psychological Bulletin*, 103(3): 345–366.

40. Nolen-Hoeksema, S., B.E. Wisco, S. Lyubomirsky. 2008. "Rethinking Rumination," *Perspectives on Psychological Science*, 3(5): 400–424.

41. Ortoleva, P. 2012. "Modeling the Change of Paradigm: Non-Bayesian Reaction to Unexpected News," *American Economic Review* 102(6): 2410–2436.

42. Ortoleva, P. 2013. "The price of flexibility: Towards a theory of Thinking Aversion," *Journal of Economic Theory* 148: 903–934.

43. Rubinstein, A. and Y. Salant. 2008. "Some Thoughts on the Principle of Revealed Preference," in: *The Foundations of Positive and Normative Economics: A HandBook*, Chapter 5, A. Caplin and A. Schotter (Eds.), Oxford University Press, Oxford.

44. Padoa-Schioppa, C. and A. Rustichini. 2014. "Rational Attention and Adaptive Coding: A Puzzle and a Solution," *American Economic Review: Papers & Proceedings*, 104(5): 507-513.

45. Rustichini, A. and Padoa-Schioppa, C. 2015. "A neuro-computational model of economic decisions," *Journal of Neurophysiology*, 114: 1382-1398.

46. Schipper, B. 2015. "Awareness," in: *Handbook of Epistemic Logic*, Chapter 3, H. van Ditmarsch, J.Y. Halpern, W. van der Hoek and B. Kooi (Eds.), College Publ., London.

47. Tversky, A. 1972. "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79: 281–99.

Cost-Benefit Analysis in Reasoning: Online Appendix

1 Appendix on the lottery elicitation method

Here we formalize the discussion in Section 4.2.1 concerning the lottery method elicitation. We first expand the mental preference relation relation to an ex-ante stage that includes lotteries. Formally, let the enriched preferences $\hat{\succsim}_s$ be a binary relation over $\mathcal{L} := \Delta^b(\mathcal{U} \times \{0,1\})$ with representative element $L, L' \in \mathcal{L}$, where $\Delta^b(\mathcal{U} \times \{0,1\})$ is the set of binary lotteries on $\mathcal{U} \times \{0,1\}$.

In any mental state s, a lottery $L' = ((1-\epsilon; (u,x)); (\epsilon; (u',y))) \in \mathcal{L}$, with $x, y \in \{0,1\}$, is interpreted in the following way. If $x = 1$ and $y = 0$, L' is a lottery that gives with probability $1-\epsilon$ payoff function u for which the agent then has to think an extra step, and with probability ϵ it gives u' for which the agent does not think further (and for which the action a^s has already been locked in). If $x = y = 1$, , we assume that the agent has to commit to reason *before* knowing whether he receives u or u'.

To express that the agent may have to think an extra step before knowing whether he receives u or u', we enrich the definition of a choice problem to include lotteries over u. Specifically, let $\Delta^b(\mathcal{U})$ be the space of lotteries over payoff functions, with typical element $P = ((1-\epsilon; u); (\epsilon; u')))$. Instead of using choice problems to be (u, E) as in the text, we will define an enriched-choice problem $(P, E) \in \Delta^b(\mathcal{U}) \times \mathcal{E}$, which is identical to a choice problem except that it allows for the reasoning to be over lotteries over u. Note that our setting of interest in the text, (u, E), corresponds to the case in which the lotteries P are degenerate; we will denote these δ_u, and refer to problems (δ_u, E) as choice problems, to emphasize that they correspond to choice problems (u, E) as defined in the text.

An enriched-mental state s comprises of an action $a^s \in a$ and preference relation $\hat{\succsim}_s$. The idea of cognitive equivalence, over this enriched space will be defined in an analogous way to Definition 1:

Definition 1A (Enriched Cognitive Equivalence)**.** *Two mental states $s, s' \in S$ are* enriched-cognitively equivalent (e.c.e.) *if $a^s = a^{s'}$ and $\hat{\succsim}_s = \hat{\succsim}_{s'}$. Two choice problems (P, E) and (P', E') are* enriched-cognitively equivalent *if they induce sequences of pairwise e.c.e. states.*

Take any choice problem (δ_u, E) and suppose that the agent reaches mental state $s \in S(\delta_u, E)$. Suppose that the agent can be asked to express preferences $\hat{\succsim}_s$ about some lotteries $L, L' \in \mathcal{L}$. Here, we will *not* presuppose that the agent uses the same thought process to reason about L or L' as he does to think about u. That is, we do not follow the thought exercise of asking the agent to think about L and L' but from the mental state associated with (δ_u, E) – which would be the analogy of the thought exercise followed in the text with $\succsim_s$. But our lottery method, if our assumptions below hold, will serve to circumvent this issue for the kinds of axioms we are interested in in the text. In particular, our assumptions will allow us to define mental preferences $\succsim_s$ to effectively capture the main text's thought exercise.

We are now in position to present our key assumptions for the appropriate use of the lottery elicitation method. Take any choice problem (δ_u, E), and now consider an enriched-choice problem $(P'^b(\mathcal{U}) \times \mathcal{E}$, where $P' = ((1-\epsilon; u); (\epsilon; u'))$ for some u'. For sufficiently small ϵ, we expect that the agent would stick to reasoning about (P', E) in the same way as he would reason about the original choice problem (δ_u, E), meaning that they are *enriched-cognitively equivalent* (which is essentially what is assumed in standard experimental practices). This logic leads to the following assumption.

Assumption 1A (Local stickiness)**.** *For any (δ_u, E) and u', where $u, u' \in \mathcal{U}$, there is a $\bar{\epsilon}(u, u') \in (0,1)$ such that, for any $\epsilon \in (0, \bar{\epsilon}(u, u'))$ and $P' = ((1-\epsilon; u); (\epsilon; u'))$, (P', E) is enriched-cognitively equivalent to (δ_u, E).*

Our second assumption concerns enriched preferences $\hat{\succsim}_s$. Consider again a choice problem (δ_u, E), and suppose now that we compare lotteries L' and L'', where $L' = ((1-\epsilon; (u, 1)); (\epsilon; (u', 1)))$ and $L'' = ((1-\epsilon; (u, 1)); (\epsilon; (u'', 1)))$, for a sufficiently small $\epsilon > 0$. Suppose, moreover, that we elicit that $L' \hat{\succsim}_s L''$. We would then expect that for another sufficiently small $\tilde{\epsilon}$ and $\tilde{L}' = ((1-\tilde{\epsilon}; (u, 1)); (\tilde{\epsilon}; (u', 1)))$ and $\tilde{L}'' = ((1-\tilde{\epsilon}; (u, 1)), (\tilde{\epsilon}; (u'', 1)))$, the same preference rankings hold, i.e., $\tilde{L}' \hat{\succsim}_s \tilde{L}''$.

Assumption 2A (Local conditional invariance)**.** *For any (δ_u, E) and u', u'', where $u, u', u'' \in \mathcal{U}$, and for all $s \in s(\delta_u, E)$ there is a $\breve{\epsilon}(u, u', u'')$ such that, for all $\epsilon, \tilde{\epsilon} \in (0, \breve{\epsilon}(u, u', u''))$, $x, y \in \{0, 1\}$ and for all $L' = ((1-\epsilon; (u, 1); (\epsilon; (u', x)))$, $L'' = ((1-\epsilon; (u, 1); (\epsilon; (u'', y)))$, $\tilde{L}' = ((1-\tilde{\epsilon}; (u, 1)); (\tilde{\epsilon}; (u', x)))$ and $\tilde{L}'' = ((1-\tilde{\epsilon}; (u, 1)); (\tilde{\epsilon}; (u'', y)))$, the following holds:*

$$L' \hat{\succsim}_s L'' \Leftrightarrow \tilde{L}' \hat{\succsim}_s \tilde{L}''.$$

We now use $\hat{\succsim}_s$ to define preferences $\succsim_s$, which are the objects used in the text, with the understanding that an enriched mental state $s \in S(\delta_u, E)$ corresponds to a (non-enriched) mental state $s \in S(u, E)$ in the model of the main text. Under the two assumptions above, the following is well-defined:

Definition 2A (Preferences $\succsim_s$)**.** *For any enriched-choice problem (δ_u, E) and enriched-mental state $s \in S(\delta_u, E)$, and corresponding (non-enriched) choice problem (u, E), preferences $\succsim_s$ are defined to be, for any $u', u'' \in \mathcal{U}$, $x, y \in \{0, 1\}$: $(u', x) \succsim_s (u'', y)$ if and only if $L' \hat{\succsim}_s L''$, where $L' = ((1-\epsilon; (u, 1)); (\epsilon; (u', x)))$, $L'' = ((1-\epsilon; (u, 1)); (\epsilon; (u'', y)))$ for some $\epsilon \in (0, \min\{\hat{\epsilon}(u, u'), \breve{\epsilon}(u, u', u'')\})$.*

In words, whenever we wish to compare $u' \succsim_s u''$ for some mental state $s \in S(u, E)$, we can use the lottery-elicitation method of comparing L' to L'' instead, for some sufficiently small probability ϵ.[1]

[1] An assumption that is closer in spirit and often made implicitly in experimental settings would be

Given this definition, all of the axioms provided in the text could be expressed in the space discussed here, and the extended relations can be elicited in practice. We keep the axioms as they are in the text as they are simpler to present and interpret.

that $((1;(u',x)))\hat{\succsim}_s((1;(u'',y))) \Leftrightarrow L'\hat{\succsim}_s L''$, meaning that a small probability of receiving u does not affect the preference rankings between (u',x) and (u'',y). Such an assumption is stronger than local conditional invariance, under the maintained assumptions (transitivity and local stickiness).