A robust measure of complexity

Egor Bronnikov[*]   Elias Tsakas[†]

Maastricht University

January, 2025

Preliminary draft

Latest version of the paper.

Abstract

We introduce a robust belief-based measure of complexity. The idea is that task A is deemed more complex than task B if the probability of solving A correctly is smaller than the probability of solving B correctly regardless of the reward. We fully characterize the corresponding order over the set of tasks. The main characteristic of this relation is that it depends, not only on difficulty (like most complexity definitions in the literature) but also on ex ante uncertainty. Finally, we show that for every task for which information is optimally acquired, there exists a more complex task which always induces less effort regardless of the reward.

Keywords: complexity, measure, difficulty, ex ante uncertainty, incomplete relation, effort.
JEL codes: D83, D90.

1.   Introduction

Complexity is a fundamental concept across numerous scientific domains, including computer science, cognitive sciences, neuroscience, etc. More recently, its importance has also been recognized by (behavioral) economists, as it is a key determinant of decisions in many settings, and more importantly it has the potential to explain mistakes that people systematically make in such decisions (Banovetz and Oprea, 2023; Enke et al., 2024a; Oprea, 2024b).

---

[*]Homepage: egorbronnikov.github.io; E-mail: egor.bronnikov@maastrichtuniversity.nl
[†]Homepage: www.elias-tsakas.com; E-mail: e.tsakas@maastrichtuniversity.nl

At the same time, although most notions of complexity somehow reflect the difficulty to handle a task —or the associated cost— it is also the case that complexity is a term which is traditionally used in a casual way, without consensus on what exactly it means (Gabaix and Graeber, 2024). This lack of a widely accepted precise definition can possibly explain why scholars have also focused on measuring complexity, which can be in turn serve as a proxy for different types of complexity (Oprea, 2024a).[1] Common ways to measure complexity include direct metrics (Oprea, 2020), behavioral metrics (Banovetz and Oprea, 2023), and most importantly for this paper belief-based metrics (Enke and Graeber, 2023; Enke et al., 2024a,b).

The key idea within this last stream of literature is to use an agent's belief about their own accuracy as a proxy for complexity. This is also consistent with theoretical results which show that people are more accurate when solving simpler tasks (Goncalves, 2024). However, as appealing as this approach is, there is a serious caveat, which stems from the fact expected accuracy depends both on complexity as well as on the effort to handle the task. And since effort depends non-linearly on the underlying reward (for correctly solving the task), it often happens that expected accuracy is sensitive to the reward. And this naturally gives rise to the question: if different rewards yield different revealed complexity order over the set of tasks, which one should we use?

In this paper, we take a robust approach, proposing the following way to rank tasks: we say that task A is revealed to be more complex than task B, whenever the chances to correctly solve A are smaller than the chances to correctly solve B, *for every (extrinsic) reward*. Obviously, this is a rather conservative criterion, as it imposes a rather strong dominance condition. But at the same time, whenever satisfied, it provides quite strong and convincing evidence that the tasks are indeed ranked in this way, thus indirectly providing a sufficient condition that every reasonable definition of complexity should satisfy.

The main result of the paper provides a full characterization of the complexity order that the aforementioned criterion induces (Theorem 1). Not surprisingly, the complexity order is incomplete. But they key insight is that it depends on two distinct parameters, viz., the difficulty of the task and the ex ante uncertainty. More specifically, higher difficulty is only a necessary condition for higher complexity. In order to also become sufficient, the agent cannot be ex ante much more uncertain about the state realization. For example, in order for a student to find exam A more complex than exam B, it must be both the case that A is more difficult, and moreover that the student is not much better prepared about B. This is a novel channel of complexity, as in most of the existing literature, complexity is typically associated only with some sort of difficulty.

---

[1]Measurements and formal definitions of complexity are closely linked with each other, in an analogous way belief elicitation mechanisms (Brier, 1950; Savage, 1971) are linked with formal notions of subjective probability became established (Savage, 1954; Anscombe and Aumann, 1963).

The fact that complexity has these two distinct dimensions allows us to establish a link with the decision-theoretic literature on incomplete preferences. In particular, we show that our complexity order is represented by a vector (utility) function, like in Ok (2002).

This representation allows us to do comparative statics with respect to (unobservable) intrinsic incentives that the agent may have on top of the extrinsic reward. In particular, we show that increasing the intrinsic incentives has a dual effect, viz., on the one hand, it leads to a more complete order (i.e., we can compare more tasks with one another), but at the same time, it also makes the differences between tasks less salient (i.e., it is less clearcut which task is more complex). In this sense,

Finally, we go back to the relationship between complexity and induced effort, where effort is seen as a proxy for the cost that the agent optimally incurs when handling the task. As it has been already pointed out in the literature, we should not expect effort to be monotonic with respect to complexity, i.e., it might be the case that simpler tasks induce more effort, as the agent may not find it worthwhile to try at all in order to answer a task which is anyway too difficult (Goncalves, 2024). However, this argument is typically obtained for fixed rewards. Thus, we naturally ask whether it is robust across different rewards.

Our second main result shows that this is indeed the case, i.e., for every task against which the agent optimally acquires some information, there exists some more complex task which always induces less effort, regardless of the size of the reward (Theorem 2). The result is quite surprising, as one would expect that for sufficiently high reward, the agent would put maximum effort both for the complex and the simple task. However, what happens here is that the more complex task that we find is a but more difficult, and crucially involves much more ex ante uncertainty. As a result, the role of increased difficulty is relatively small.

The literature on complexity is vast, and as such we are de facto forced to make a selection of what in our view is the most relevant subset.

Early work focused primarily on the role of strategy complexity within game theory (Rubinstein, 1986; Abreu and Rubinstein, 1988). More recently, the focus has shifted towards explaining mistakes and irrationalities, e.g., Oprea (2024b) study the effect of complexity on risk preferences, and Enke et al. (2024a) the respective effect on time preferences. There is also interest in formalizing definitions of complexity, e.g., Gabaix and Graeber (2024) build a general model of production within a cognitive economy in order to operationalize complexity, whereas Oprea (2024a) borrows insights from computer science to introduce a framework within which complexity reflects the cost for handling a task. Others, define it as the signal-to-noise ratio (Goncalves, 2024), similarly what is often done in psychometrics. The common denominator throughout most of this literature is that definitions of complexity are typically input-based, i.e., they somehow reflect the underlying difficulty to process and handle a task.

What is closer to our work is the literature on measuring complexity. As Oprea (2024a)

elegantly points out, this literature can be classified into three large streams, depending the measurement tool. Within the first stream, we encounter measurement with direct metrics, such as willingness to pay in order to avoid dealing with a certain task (Oprea, 2020), response times (Goncalves, 2024), and biometrics (van der Wel and van Steenbergen, 2018). The second stream leverages behavioral metrics, such as procedural measurements (Banovetz and Oprea, 2023), and choice inconsistencies (Woodford, 2020).

Finally, the third stream, within which our paper belongs, uses belief-based metrics. These include subjective rankings, like for instance in Gabaix and Graeber (2024) where subjects are simply asked to rank tasks with respect to complexity, and beliefs about optimality of the subject's own accuracy (Enke and Graeber, 2023; Enke et al., 2024a,b), like in our paper. Of course, in all aforementioned papers, rewards are fixed. The effects of varying rewards are discussed in (Alaoui and Penta, 2022).

This entire literature is part of a surging field of Cognitive Economics (Caplin, 2025; Enke, 2024), which also incorporates topics such as rational inattention, cognitive uncertainty, etc.

The paper is structured as follows: In Section 2 we introduce our measure of complexity and prove our main characterization result. In Section 3 we study properties of the complexity order which is induced by our measure. In Section 4 we study the relationship between our complexity measure and the induced effort.

2. A measure of subjective complexity

Take a binary state space $\Theta = \{\theta_0, \theta_1\}$. An agent is asked to guess the realization of Θ. Let $Y = \{y_0, y_1\}$ denote the possible results of this guessing task, i.e., y_0 denotes a wrong guess and y_1 denotes a correct guess. Let $X := [x_0, \infty)$ be a convex set of monetary rewards for guessing correctly. In this sense, Y is the set of intrinsic rewards, and X is the set of extrinsic rewards.

The agent's preferences over the set of acts $(X \times Y)^\Theta$ admit a SEU representation with Bernoulli utility function $u : X \times Y \to \mathbb{R}$, which is often abbreviated by $u_0(x) := u(x, y_0)$ and $u_1(x) := u(x, y_1)$. For both $y \in Y$, the utility function is continuously increasing and unbounded in X. We use the difference

$$w := u_1(x_0) - u_0(x_0) \tag{1}$$

as a measure of intrinsic motivation, i.e., this is how much the agent cares about being right, in the absence of extrinsic rewards. We assume that $w \geq 0$. Note that the previous inequality is only weak, i.e., in the absence of extrinsic incentives, the agent may or may not care about intrinsic incentives. Furthermore, observe that we do not impose any condition on the difference $u_1(x) - u_0(x)$ for any $x > x_0$. Neither do we assume separability of u with respect to X and Y. Finally, without loss of generality, the utility function is normalized so that $u(x_0, y_0) = 0$.

Going back to the guessing task, for some fixed extrinsic reward $x \in X$, the guess $r \in \Theta$ is an act which yields utility $u_1(x)$ at state $\theta = r$ and utility 0 at $\theta \neq r$. Thus, the agent's indirect expected utility, as a function of the probability $q \in [0, 1]$ that she attaches to θ_1, is given by

$$g(q) := u_1(x) \max\{q, 1 - q\}, \tag{2}$$

which is obviously proportional to the probability she attaches to her best guess being correct.

Suppose that the agent holds a prior belief which assigns probability $p \in [0, 1]$ to θ_1. This belief incorporates her prior knowledge/experience, and therefore represents the degree of her ex ante uncertainty in terms of proximity to the (maximally uncertain) uniform belief:

$$\phi := 1 - 2|p - 1/2|. \tag{3}$$

That is, the larger ϕ, the more ex ante uncertain the agent is, with $\phi = 0$ meaning that the prior attaches probability 1 to one of the two states, and $\phi = 1$ meaning that p is uniformly distributed. This notion of uncertainty resembles the usual measures of uncertainty from information theory (Cover and Thomas, 2006). Then, before making a guess, the agent decides how much attention to pay. Attention is modelled with a Bayesian signal, which is uniquely identified by a mean-preserving distribution of posterior probabilities (Kamenica and Gentzkow, 2011). The set of all signals is denoted by

$$\Pi(p) = \left\{ \pi \in \Delta\big([0, 1]\big) : \mathbb{E}_\pi(q) = p \right\}.$$

For any signal $\pi \in \Pi(p)$, define the agent's ex ante indirect expected utility,

$$G(\pi) := \mathbb{E}_\pi\big(g(q)\big). \tag{4}$$

It is not difficult to see that $G(\pi)$ is proportional to the probability $G(\pi)/u_1(x)$ that she attaches ex ante (i.e., before π is realized) to her best guess being eventually correct.

Aligned with the rational inattention literature, we assume that attention is costly. The cost function is assumed to be uniformly posterior separable (Caplin et al., 2022), i.e., there is a strictly convex function $c : [0, 1] \to \mathbb{R}$ such that the cost of signal $\pi \in \Pi(p)$ is given by

$$C(\pi) = \kappa\Big(\mathbb{E}_\pi\big(c(q)\big) - c(p)\Big), \tag{5}$$

where $c(q)$ represents the agent's marginal cost for acquiring information, and $\kappa > 0$ is a parameter of the task's difficulty. Note that consistently with the complexity literature (Oprea, 2024a), the cost consists an objective part (viz., the parameter κ) and a subjective part (viz., the function c). Such cost functions have solid foundations (Denti, 2022) and are supported by experimental evidence (Dean and Neligh, 2023). Throughout the paper, we will focus on

symmetric cost functions, which include the Shannon entropy (Sims, 2003), the Shorrocks entropy (Shorrocks, 1980), the Tsallis entropy (Caplin et al., 2022) as special cases. Formally, this means that for every $q \in [0,1]$, we have $c(q) = c(1-q)$. For an axiomatization of symmetric cost functions, see Hébert and Woodford (2021).

The agent faces a tradeoff, in that more informative signals help her to achieve higher expected utility, but at the same time are also more costly. That is, formally speaking, the agent solves the following optimization problem:

$$\max_{\pi \in \Pi(p)} \Big(G(\pi) - C(\pi) \Big). \tag{6}$$

In the previous optimization problem, four parameters enter the picture: the extrinsic reward x, the intrinsic incentives w, the ex ante uncertainty ϕ, and the difficulty κ. The only one which is usually observable by the analyst and can potentially vary is the extrinsic reward x. The intrinsic reward w is typically unobservable, but most importantly cannot vary, as it simply reflects the agent's exogenous satisfaction for solving a task correctly. In this sense, throughout the paper, it is treated as an individual characteristic of the agent. Finally, the ex ante uncertainty ϕ and the difficulty κ are inherently linked with the task itself, and in particular they represent the past information and the cost for acquiring new information, respectively. In this sense, we will henceforth treat the pair (ϕ, κ) as a description of the task, and correspondingly

$$\mathcal{T} := [0,1] \times (0, \infty)$$

as the set of all tasks.

It is not difficult to verify that for every tuple of parameters (x, w, ϕ, κ) there is a unique solution to the optimization problem (6), henceforth denoted by $\pi(\cdot|x, w, \phi, \kappa)$. It follows from standard arguments (e.g., Kamenica and Gentzkow, 2011; Matějka and McKay, 2015) that the optimal signal is given by concavifying the function $g(q) - \kappa c(q)$, as illustrated in the figure below.

Then, the agent's subjective expected accuracy is given by

$$F(x, w, \phi, \kappa) := \frac{G\big(\pi(\cdot|x, w, \phi, \kappa)\big)}{u_1(x)}. \tag{7}$$

This is equal to the expected probability of guessing correctly given that the optimal signal is used $\pi(\cdot|x, w, \phi, \kappa)$.

Now, consistently with the existing literature which suggests that complexity is positively correlated with accuracy (Goncalves, 2024), we would like to use $F(x, w, \phi, \kappa)$ as a measure of subjective complexity. However, there is still a caveat: In our view, a sensible definition of complexity should be disentangled from the material reward that the agent will receive in case
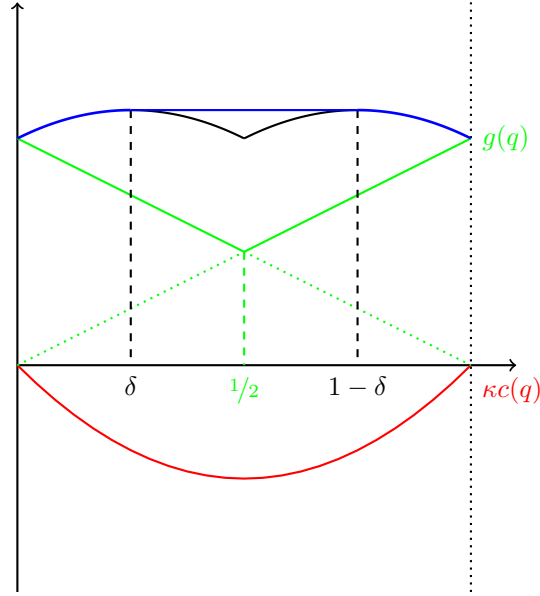
Figure 1: The green solid piecewise linear function is the net expected utility from guessing correctly, and the red line is the marginal cost for acquiring information. The blue curve is the concave closure of the difference $g(q) - \kappa c(q)$. The optimal signal distributes all the probability between the posteriors δ and $1 - \delta$ whenever ex ante uncertainty is large (i.e., whenever $p \in (\delta, 1 - \delta)$), and it is completely uninformative otherwise.
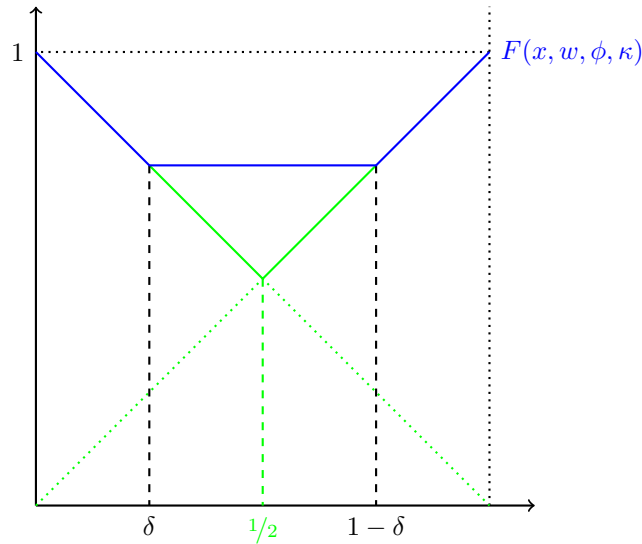

Figure 2: The blue piecewise linear function is the expected accuracy, assuming that the optimal signal has been used.

she solves the task. The analogy is simple. Subjective complexity should only be linked with characetristics of the task itself. In this sense, we will use subjective expected accuracy as a

measure of subjective complexity only when it is robust with respect to the extrinsic reward, i.e., formally, we introduce the following definition.

Definition 1. For an agent whose intrinsic incentives are given by $w \geq 0$, we say that task (ϕ, κ) is **subjectively more complex** than task (ϕ', κ'), and we write $(\phi, \kappa) \succeq_w (\phi', \kappa')$, whenever

$$F(x, w, \phi, \kappa) \leq F(x, w, \phi', \kappa') \tag{8}$$

for all $x \in X$. ◁

The asymmetric and the symmetric parts of $\succeq_w$ are defined as usual. That is, we have $(\phi, \kappa) \succ_w (\phi', \kappa')$ whenever $(\phi, \kappa) \succeq_w (\phi', \kappa')$ and $(\phi, \kappa) \not\preceq_w (\phi', \kappa')$, and respectively $(\phi, \kappa) \sim_w (\phi', \kappa')$ whenever $(\phi, \kappa) \succeq_w (\phi', \kappa')$ and $(\phi, \kappa) \preceq_w (\phi', \kappa')$.

A task (ϕ, κ) is said to be trivial given intrinsic incentives w, if the optimal signal reveals the true state with certainty for every $x \geq x_0$, i.e., the support of the optimal signal $\pi(\cdot | x, w, \phi, \kappa)$ contains only posteriors that put probability 1 to a single state in Θ. Obviously, if a task (ϕ, κ) is trivial, then it is simpler than any other task, i.e., $(\phi', \kappa') \succeq_w (\phi, \kappa)$ for all $(\phi', \kappa') \in \mathcal{T}$. The set of non-trivial tasks is henceforth denoted by $\mathcal{T}_w \subseteq \mathcal{T}$, and it is characterized by a difficulty threshold $\kappa_w \geq 0$.

Proposition 1. *For each $w \geq 0$ there is some $\kappa_w \geq 0$, such that the following are equivalent:*

(i) $\kappa > \kappa_w$.

(ii) $(\phi, \kappa) \in \mathcal{T}_w$ for every $\phi \in [0, 1]$.

Moreover, κ_w is increasing in w.

The idea is quite simple: in order for a task to be non-trivial, the information costs must be sufficiently large to guarantee that the intrinsic incentives alone are not strong enough to always lead to a perfectly informative signal. In our main result below, we characterize how non-trivial tasks are ranked in terms of subjective complexity.

Theorem 1. *For every fixed $w \geq 0$, there is an increasing function $\phi_w : (\kappa_w, \infty) \to [0, 1]$ such that the following are equivalent for any pair (ϕ, κ) and (ϕ', κ') of non-trivial tasks:*

(i) $(\phi', \kappa') \succeq_w (\phi, \kappa)$.

(ii) $\kappa' \geq \kappa$ and $\phi' \geq \min\{\phi_w(\kappa), \phi\}$.

8

Let us first explain what $\phi_w(\kappa)$ stands for. Suppose that there is no extrinsic reward, i.e., let $x = x_0$. Then, for a given level of intrinsic incentives $w \geq 0$, the agent will acquire information if and only if her ex ante uncertainty satisfies the following inequality:

$$\phi > \phi_w(\kappa), \tag{9}$$

i.e., the prior needs to lie sufficiently close to the uniform belief in order for the agent to optimally acquire information. Obviously, ϕ_w is continuously decreasing with $\phi_w(\kappa) \to 0$ as $\kappa \to \kappa_w$, and $\phi_w(\kappa) \to 1$ as κ grows arbitrarily large. This means without sufficiently large costs, the agent will always acquire the perfectly informative signal regardless of her prior, whereas when the cost becomes infinitely large she will not acquire information regardless of her prior belief. Moreover, ϕ_w is increasing in w, i.e., the stronger the intrinsic incentives are, the easier it becomes to acquire information.

Graphically, the previous result is illustrated in the Figures below: which correspond to two cases, i.e., the one where the information-acquisition constraint of (9) holds, and the one where it does not. In the figures, we have taken $c(q) = q^2 - q$, which is subdifferentiable at the boundaries of the unit interval, and therefore for small cost parameters the agent will optimally acquire the perfectly informative signal regardless of the size extrinsic reward. This explains why ϕ_w is initially constant at 0, and as a result there is a grey region which contains the trivial tasks. Such horizontal part would not have appeared if the cost function was for instance entropic, in which case a perfectly informative signal is never optimal.

Starting with the first case (Figure 3), we take a task (ϕ, κ) such that $\phi > \phi_w(\kappa)$. Then, the red region contains the tasks (ϕ', κ') that are subjectively more complex than (ϕ, κ). On the other hand, the green region contains the tasks (ϕ', κ') that are subjectively simpler than (ϕ, κ). How we obtained the red region is obvious given our previous theorem. So, let us elaborate on how the green region arises. First of all, $\kappa \geq \kappa'$ and $\phi \geq \phi'$ jointly imply $(\phi, \kappa) \succeq_w (\phi', \kappa')$, again by our theorem. Then, let us consider some (ϕ', κ') such that $\kappa' < \kappa$ and $\phi' \geq \phi$. By ϕ_w being increasing, it follows that $\phi_w(\kappa') < \phi_w(\kappa) < \phi$. Hence, once again by our previous theorem, we obtain $(\phi, \kappa) \succeq_w (\phi', \kappa')$.

Let us now focus on the second case (Figure 4), where we take a task (ϕ, κ) such that $\phi \leq \phi_w(\kappa)$. Once again, by our previous theorem, the red region is the set of tasks that are subjectively more complex than (ϕ, κ). And once again, the green region is the set of tasks that are subjectively simpler than (ϕ, κ). Regarding this last part, let us elaborate only on the non-obvious case where $\kappa' < \kappa$ and $\phi' > \phi$. If κ' is such that $\phi_w(\kappa') > \phi$, then by our theorem (ϕ, κ) and (ϕ', κ') are not comparable via $\succeq_w$.

Remark 1. Note that in most of the existing literature, complexity is taken as a synonym of difficulty. Here we show that this is not the case, as complexity depends both on difficulty and ex ante uncertainty. This is consistent with the idea that a task may be deemed complex
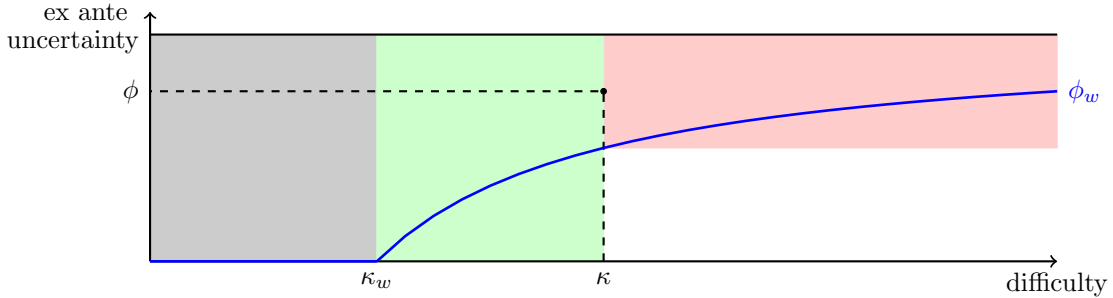
Figure 3: This is the case when the ex ante uncertainty is large and therefore it is optimal to acquire information before making a guess. The red area contains the tasks that are deemed more complex than (ϕ, κ), and the green area are the tasks that are deemed simpler than (ϕ, κ). Finally, the grey area contains the trivial tasks, meaning that they are deemed simpler than every other task, including (ϕ, κ).
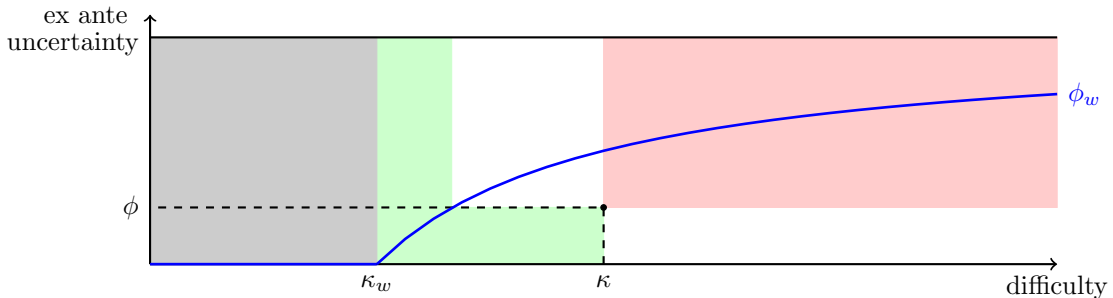

Figure 4: This is the case when the ex ante uncertainty is small and therefore it is optimal not to acquire any information before making a guess. Once again, the red area contains the tasks that are deemed more complex than (ϕ, κ), the green area the tasks that are deemed simpler than (ϕ, κ), and the grey area the trivial tasks.

because it has not been previously encountered by the agent. For instance, students often complain that they did not have enough practice questions similar to the ones they faced in their final exam, even though their actual exam was not particularly difficult. Nevertheless, even though increased difficulty is not sufficient for increased complexity, it is still a necessary condition, i.e., unless both (ϕ, κ) and (ϕ', κ') are both trivial, $(\phi', \kappa') \succeq_w (\phi, \kappa)$ will necessarily imply $\kappa' \geq \kappa$. ◁

3. Properties of subjective complexity order

Our definition of subjective complexity is robust with respect to the extrinsic reward. However, such robustness comes at the cost of not being able to rank all tasks. Let us provide some results on the structure of $\succeq_w$ for different $w \geq 0$.

Proposition 2. *For every $w \geq 0$, the relation $\succeq_w$ is transitive and incomplete.*

The previous result can be easily illustrated in the two figures above, i.e., there is always a white region which includes the tasks that are subjectively neither more complex nor simpler than any given task.

Note that within $\mathcal{T}_w$ the relation $\succeq_w$ has the same structure as the incomplete preference relations in Ok (2002). To understand the connection, recall that in that paper preferences are represented by a vector-valued utility function. In the context of our paper this would mean that there exists some

$$\boldsymbol{u} : \mathcal{T}_w \to \mathbb{R}^2$$

such that $(\phi, \kappa) \succeq_w (\phi', \kappa')$ if and only if $\boldsymbol{u}_1(\phi, \kappa) \geq \boldsymbol{u}_1(\phi', \kappa')$ and $\boldsymbol{u}_2(\phi, \kappa) \geq \boldsymbol{u}_2(\phi', \kappa')$, where $\boldsymbol{u}_1(\phi, \kappa) := \kappa$ and $\boldsymbol{u}_2(\phi, \kappa) := \min\{\phi_w(\kappa), \phi\}$. The latter also distinguishes our notion of complexity from other definitions in the literature which typically induce a complete order, e.g., the signal-to-noise ratio (Callander, 2011; Fehr and Rangel, 2011; Goncalves, 2024).

Now, let us look at the comparative statics with respect to the intrinsic incentives. This will provide us some insight on whether we should be using more or less intrinsically motivated agents to compare tasks in terms of complexity.

Proposition 3. *If $w' > w$, then the following hold:*

(a) $\mathcal{T}_{w'} \subseteq \mathcal{T}_w$.

(b) $\succeq_w \subseteq \succeq_{w'}$.

Remark 2. Monotonicity of $\succeq_w$ does not extend to monotonicity of the asymmetric part $\succ_w$ of the relation. Indeed, it might be the case that $(\phi, \kappa) \succ_w (\phi', \kappa')$ for some w, and then $(\phi, \kappa) \sim_{w'} (\phi', \kappa')$ for some $w' > w$. This is, for instance, the case when $\kappa = \kappa'$ and $\phi' > \phi_w(\kappa) > \phi > \phi_{w'}(\kappa)$. However, what will never happen is a complete reversal of subjective complexity, i.e., it cannot be that $(\phi, \kappa) \succ_w (\phi', \kappa')$ and $(\phi, \kappa) \prec_{w'} (\phi', \kappa')$. ◁

Going back to Proposition 3, the fact that $\succeq_w$ is increasing in w follows from ϕ_w being pointwise decreasing, i.e., if $w' > w$, then $\phi_{w'}(\kappa) \leq \phi_w(\kappa)$ for every $\kappa > 0$. The intuition is pretty clear: stronger intrinsic incentives will lead to more frequent information acquisition. At the same time, more tasks will become trivial, and therefore equivalent in terms of complexity. In fact, it is not difficult to see that in the limit, as w becomes arbitrarily large, agent deems all tasks equally complex.

Together, Proposition 3 and Remark 2, point to an interesting tradeoff. Stronger intrinsic incentives allow us to compare more tasks in terms of complexity, but the differences between tasks become less salient. This is because expected accuracy becomes inflated by the intrinsic incentives, and therefore all expected accuracies are concentrated near extreme beliefs, which makes it more difficult to distinguish the complexity of different tasks.

4. Subjective complexity and effort

The relationship between complexity and effort has been studied by several papers in the literature (Goncalves, 2024, and references therein). In this paper, we naturally assume that expected response time is a proxy for effort, which is in turn positively correlated with the cost $C\big(\pi(\cdot|x, w, \phi, \kappa)\big)$ that the agent incurs for optimally acquiring information about the task.

For starters, it is not difficult to see that for any fixed $x \geq x_0$, the induced effort is not always monotonic in complexity. This finding is not surprising, and it has been already pointed out in the literature. Here we will focus on the the problem from a different angle, focusing on robustness with respect to the extrinsic reward, i.e., we ask whether the non-monotonicity can arise for every $x \geq x_0$.

Theorem 2. *Fix $w \geq 0$, and take an arbitrary $(\phi, \kappa) \in \mathcal{T}_w$ such that $\phi > \phi_w(\kappa)$. Then, there exists $(\phi', \kappa') \succ_w (\phi, \kappa)$ such that $C\big(\pi(\cdot|x, w, \phi', \kappa')\big) < C\big(\pi(\cdot|x, w, \phi, \kappa)\big)$ for all $x \geq x_0$.*

The previous result shows that for all tasks for which the agent optimally acquires some information, we can always find a more complex task which will induce less effort regardless of the size of the extrinsic reward. In this sense, non-monotonicity of effort with respect to complexity is both generic and robust with respect to the extrinsic incentives. Graphically this is illustrated in Figure 5 below. Of course, for extrinsic reward x_0 it is pretty obvious, as the agent will not acquire any information when facing (ϕ', κ'), as opposed to when facing (ϕ, κ). The interesting part arises when the extrinsic reward increases to any $x > x_0$. In this case, although (ϕ', κ') is harder, it involves lower ex ante uncertainty. And this is exactly what makes it cheaper.
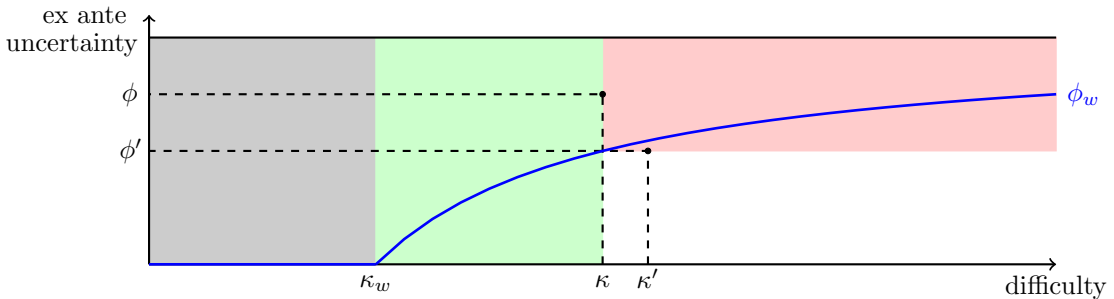

Figure 5: THere is some $\kappa' > \kappa$ such that task (ϕ', κ') is strictly more complex than (ϕ, κ), but at the same time it always induces less effort, regardless of how large the extrinsic reward is.

In fact, it is not difficult to show that if complexity is entirely driven by ex ante uncertainty, more complex tasks will always induce more effort. That is, whenever $\phi > \phi'$, it will be the case that $(\phi, \kappa) \succ_w (\phi', \kappa)$, and furthermore $C\big(\pi(\cdot|x, w, \phi, \kappa)\big) > C\big(\pi(\cdot|x, w, \phi', \kappa)\big)$ for all $x \geq x_0$.

Interestingly, while the effect of ex ante uncertainty is clear, the same is not true for difficulty.

Proposition 4. *For every ϕ, there exists some $\kappa_\phi > \kappa_w$ such that, for any two $\kappa' > \kappa > \kappa_\phi$:*

$$
\begin{aligned}
C\big(\pi(\cdot|x, w, \phi, \kappa)\big) &> C\big(\pi(\cdot|x, w, \phi, \kappa')\big), \\
C\big(\pi(\cdot|x', w, \phi, \kappa)\big) &< C\big(\pi(\cdot|x', w, \phi, \kappa')\big),
\end{aligned}
$$

for extrinsic rewards $x' > x \geq x_0$.

The previous result suggests that effort rankings are never robust along the difficulty dimension. Thus, besides the fact that effort is not monotonic with respect to complexity, the strength of the extrinsic reward also matters, in the sense that for relatively small rewards the simpler task induces more effort, whereas for high rewards the more complex task induces more effort. This result can be seen as the robust counterpart of Goncalves (2024, Prop. 3), who shows that for a fixed x, effort is not monotonic with respect to difficulty.

A. Proofs

Proof of Proposition 1. Fix some $w \geq 0$. By a standard concavification argument, there exists some $\delta(x, w, \kappa) \in [0, 1/2]$ such that the optimal signal distributes all probability between $\delta(x, w, \kappa)$ and $1 - \delta(x, w, \kappa)$ whenever $p \in \big(\delta(x, w, \kappa), 1 - \delta(x, w, \kappa)\big)$, and it is completely uninformative otherwise. Note that $\delta(x, w, \kappa)$ is increasing in κ and decreasing in x. Then, define

$$
\kappa_w := \sup \big\{\kappa > 0 : \delta(x_0, w, \kappa) = 0\big\}. \tag{A.1}
$$

Note that $\kappa_w = 0$ if and only if $x_0 = w = 0$.

$(i) \Rightarrow (ii)$: By construction, for every $\kappa \leq \kappa_w$ and every $x \geq x_0$, it is the case that $\delta(x, w, \kappa) \leq \delta(x_0, w, \kappa) \leq \delta(x_0, w, \kappa_w) = 0$. This means that (ϕ, κ) is trivial given w for every $\phi \in [0, 1]$.

$(ii) \Rightarrow (i)$: Suppose that $\kappa > \kappa_w$. Then, by construction $\delta(x_0, w, \kappa) > 0$. Therefore, there is some $\phi > 0$ sufficiently close to 0 such that the optimal signal is completely uninformative. This means that $F(x_0, w, \phi, \kappa) < F(x_0, w, \phi, \kappa_w) = 1$, i.e., (ϕ, κ) is not trivial given w.

Finally, by (A.1) it follows directly that κ_w is increasing in w. $\qquad \square$

Proof of Theorem 1. Fix the intrinsic incentives $w \geq 0$, and take an arbitrary pair (ϕ, κ). Similarly to the proof of Proposition 1, for every $x \geq x_0$ there exists some $\delta_x \in [0, 1/2]$ such that the optimal signal distributes all probability between δ_x and $1 - \delta_x$ whenever $p \in (\delta_x, 1 - \delta_x)$, and it is completely uninformative otherwise. Therefore, we obtain

$$
F(x, w, \phi, \kappa) = \begin{cases} 1 - \phi/2 & \text{if } \phi \leq \phi_w^x, \\ 1 - \delta_x & \text{if } \phi > \phi_w^x, \end{cases} \tag{A.2}
$$

13

where $\phi_w^x(\kappa) := 1 - 2\left|\delta_x - 1/2\right|$ is the threshold of the degree of ex ante uncertainty for acquiring information. Then, define $\phi_w(\kappa) := \phi_w^{x_0}(\kappa)$.

$(ii) \Rightarrow (i)$: Take $\kappa' \geq \kappa$ and $\phi' \geq \min\{\phi_w(\kappa), \phi\}$, and consider two cases:

First, let $\phi' \geq \phi$. Note that Equation (A.2) can be equivalently rewritten as

$$F(x, w, \phi, \kappa) = \max\{1 - \phi/2, 1 - \delta_x\}. \tag{A.3}$$

Letting δ_x' be the low posterior we obtain from the concavification exercise given the parameters (ϕ', κ'), it is not difficult to verify that $\delta_x \leq \delta_x'$. Hence, it follows directly from (A.3) that $F(x, w, \phi', \kappa') \leq F(x, w, \phi, \kappa)$ for all $x \geq x_0$, and a fortiori we obtain $(\phi', \kappa') \succeq_w (\phi, \kappa)$.

Second, let $\phi > \phi' \geq \phi_w(\kappa)$. By definition, $\phi_w^x(\kappa)$ is decreasing in x, and hence $\phi > \phi_w^x(\kappa)$. So, combined with (A.2), we obtain $F(x, w, \phi, \kappa) = F(x, w, \phi', \kappa) = 1 - \delta_x$. Moreover, as shown in the first case above, F is decreasing in κ and therefore $F(x, w, \phi', \kappa') \leq F(x, w, \phi', \kappa)$. Combining the two yields $F(x, w, \phi', \kappa') \leq F(x, w, \phi, \kappa)$ for all $x \geq x_0$, and a fortiori we obtain $(\phi', \kappa') \succeq_w (\phi, \kappa)$.

$(i) \Rightarrow (ii)$: Suppose that $\kappa_w < \kappa' < \kappa$. Then, there exists some $x \geq x_0$ such that $\phi > \phi_w^x(\kappa)$, and therefore $F(x, w, \phi, \kappa) = 1 - \delta_x$. But then, by (A.3), it is also the case that $F(x, w, \phi', \kappa') \geq 1 - \delta_x'$. And by $\kappa' < \kappa$, we have $1 - \delta_x' > 1 - \delta_x$, and a fortiori $F(x, w, \phi', \kappa') > F(x, w, \phi, \kappa)$, meaning that $(\phi', \kappa') \not\succeq_w (\phi, \kappa)$.

Finally, suppose that $\kappa_w < \kappa \leq \kappa'$ and $\phi' < \min\{\phi, \phi_w(\kappa)\}$. The latter implies that $\phi' < \phi$. Since both (ϕ, κ) and (ϕ', κ') are non-trivial given w, it means that $F(x, w, \phi, \kappa) = 1 - \phi/2 < 1 - \phi'/2 = F(x, w, \phi, \kappa)$, and therefore $(\phi', \kappa') \not\succeq_w (\phi, \kappa)$. $\square$

Proof of Proposition 2. TRANSITIVITY: It follows directly from Theorem 1. Namely, begin with $(\phi'', \kappa'') \succeq_w (\phi', \kappa') \succeq_w (\phi, \kappa)$. Hence, we have $\kappa'' \geq \kappa' \geq \kappa$ and $\phi'' \geq \min\{\phi', \phi_w(\kappa')\}$ and $\phi' \geq \min\{\phi, \phi_w(\kappa)\}$. Then, we consider two cases:

(i) $\phi'' \geq \phi'$: It follows directly $\phi'' \geq \min\{\phi, \phi_w(\kappa)\}$, and therefore $(\phi'', \kappa'') \succeq_w (\phi, \kappa)$.

(i) $\phi' > \phi'' \geq \phi_w(\kappa')$: By monotonicity of ϕ_w we have $\phi_w(\kappa') \geq \phi_w(\kappa) \geq \min\{\phi, \phi_w(\kappa)\}$. Hence, we obtain $\phi'' \geq \min\{\phi, \phi_w(\kappa)\}$, and a fortiori $(\phi'', \kappa'') \succeq_w (\phi, \kappa)$.

INCOMPLETENESS: By the fact that ϕ_w approaches 1 in the limit, there exists some $\kappa > \kappa_w$ such that $\phi_w(\kappa) > 0$. Then, consider $\phi > \phi_w(\kappa) > \phi'$, and $\kappa' > \kappa$. By Theorem 1, we obtain $(\phi, \kappa) \not\succeq_w (\phi', \kappa')$ and $(\phi', \kappa') \not\succeq_w (\phi, \kappa)$, meaning that $\succeq_w$ is incomplete. $\square$

Proof of Proposition 3. (a) By construction, if $w' > w$ then $\phi_{w'}(\kappa) \leq \phi_w(\kappa)$ for every $\kappa > 0$. This means first of all that $\mathcal{T}_{w'} \subseteq \mathcal{T}_w$.

(b) Let $w' > w$, and take some $(\phi', \kappa') \notin \mathcal{T}_{w'}$ such that $(\phi, \kappa) \succeq_w (\phi', \kappa')$. This means that $\kappa \geq \kappa'$ and $\phi \geq \min\{\phi', \phi_w(\kappa')\}$. Hence, we also have $\phi \geq \min\{\phi', \phi_{w'}(\kappa')\}$, and therefore $(\phi, \kappa) \succeq_{w'} (\phi', \kappa')$. $\qquad\square$

Proof of Theorem 2. Let $\phi' := \phi_w(\kappa)$. Denote by p and p' two arbitrary priors that correspond to ϕ and ϕ' respectively. Then, for every $x \geq x_0$, we obtain

$$C\big(\pi(\cdot|x, w, \phi, \kappa)\big) - C\big(\pi(\cdot|x, w, \phi', \kappa)\big) = \kappa\big(c(p) - c(p')\big) = \varepsilon > 0. \tag{A.4}$$

Take some $\kappa' > \kappa$ such that $(\kappa' - \kappa)\big(c(0) - c(1/2)\big) = \varepsilon$. Then, for any $x \geq x_0$, we obtain

$$C\big(\pi(\cdot|x, w, \phi', \kappa')\big) - C\big(\pi(\cdot|x, w, \phi', \kappa)\big) = \kappa'\big(c(\delta'_x) - c(p')\big) - \kappa\big(c(\delta_x) - c(p')\big) < \varepsilon, \tag{A.5}$$

where δ_x and δ'_x are the posteriors that we obtain from the concavification exercise, like in the proof of Theorem 1. The last inequality follows from $c(\delta_x), c(\delta'_x) \leq c(0)$ and $c(p') > c(1/2)$. Hence, by (A.4) and (A.5), we obtain $C\big(\pi(\cdot|x, w, \phi', \kappa')\big) - C\big(\pi(\cdot|x, w, \phi, \kappa)\big) < 0$. $\qquad\square$

Proof of Proposition 4. Let $\phi < 1$, and define $\kappa_\phi := \phi_w^{-1}(\phi)$. Fix $\kappa' > \kappa > \kappa_\phi$, and observe that there is some $x > x_0$ such that $\pi(\cdot|x, w, \phi, \kappa')$ is completely uninformative as opposed to $\pi(\cdot|x, w, \phi, \kappa)$ which provides some information. Therefore, we obtain $C\big(\pi(\cdot|x, w, \phi, \kappa)\big) > C\big(\pi(\cdot|x, w, \phi, \kappa')\big)$. Moreover, for each $x'' \geq 0$, denote by δ_x and δ'_x the low posteriors in the support of $\pi(\cdot|x'', w, \phi, \kappa)$ and $\pi(\cdot|x'', w, \phi, \kappa')$ respectively, like in the proof of Theorem 1. Note that, as x'' becomes arbitrarily large, both δ_x and δ'_x converge to 0. Thus, by continuity of c, there exists some x' such that $C\big(\pi(\cdot|x', w, \phi, \kappa)\big) < C\big(\pi(\cdot|x', w, \phi, \kappa')\big)$. $\qquad\square$

References

Abreu, D. and Rubinstein, A. (1988). The structure of nash equilibria in repeated games with finite automata. *Econometrica*, 56:1259–1281.

Alaoui, L. and Penta, A. (2022). Cost-benefot analysis in reasoning. *Journal of Political Economy*, 130(4).

Anscombe, F. and Aumann, R. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205.

Banovetz, J. and Oprea, R. (2023). Complexity and procedural choice. *American Economic Journal: Microeconomics*, 15:3913–3951.

Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78:1–3.

Callander, S. (2011). Searching and learning by trial and error. *American Economic Review*, 111:2277–2308.

Caplin, A. (2025). *An introduction to cognitive economics: The science of mistakes.* Palgrave.

Caplin, A., Dean, M., and Leahy, J. (2022). Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Journal of Political Economy*, 130:1676–1715.

Cover, T. and Thomas, J. (2006). *Elements of Information Theory.* Wiley-Interscience.

Dean, M. and Neligh, N. (2023). Experimental tests of rational inattention. *Journal of Political Economy*, 131:3415–3461.

Denti, T. (2022). Posterior separable cost of information. *American Economic Review*, 112:3215–3259.

Enke, B. (2024). The cognitive turn in behavioral economics. *Working Paper*.

Enke, B. and Graeber, T. (2023). Cognitive uncertainty. *Quarterly Journal of Economics*, 138:2021–2067.

Enke, B., Graeber, T., and Oprea, R. (2024a). Complexity and time. *Journal of the European Economic Association*, forthcoming.

Enke, B., Graeber, T., Oprea, R., and Yang, J. (2024b). Behavioral attenuation. *Working Paper*.

Fehr, E. and Rangel, A. (2011). Neuroeconomic foundations of economic choice – recent advances. *Journal of Economic Perspectives*, 25:3–30.

Gabaix, X. and Graeber, T. (2024). The complexity of economic decisions. *Working Paper*.

Goncalves, D. (2024). Speed, accuracy, and complexity. *Working Paper*.

Hébert, B. and Woodford, M. (2021). Neighborhood-based information costs. *American Economic Review*, 111:3225–3255.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101:2590–2615.

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

Ok, E. (2002). Utility representation of an incomplete preference relation. *Journal of Economic Theory*, 104:429–449.

Oprea, R. (2020). What makes a rule complex? *American Economic Review*, 110:3913–3951.

Oprea, R. (2024a). Complexity and its measurement. *Handbook of Experimental Methods in the Social Sciences*, forthcoming.

Oprea, R. (2024b). Decisions under risk are decisions under complexity. *American Economic Review*, 114:3789–3811.

Rubinstein, A. (1986). Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, 39:83–96.

Savage, L. (1954). *The foundations of statistics*. Wiley.

Savage, L. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66:783–801.

Shorrocks, A. (1980). The class of additively decomposable inequality measures. *Econometrica*, 48:613–625.

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

van der Wel, P. and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin and Review*, 25:2005–2015.

Woodford, M. (2020). Modeling imprecision perception, valuation and choice. *Annual Review of Economics*, 12.