Expected accuracy as a measure of subjective complexity*

Egor Bronnikov† Elias Tsakas‡

Maastricht University

First draft: January 2025; This draft: February 2026

Latest version of the paper.

Abstract

We introduce uniform expected accuracy as a proxy for subjective complexity which is robust with respect to the underlying reward for solving the task correctly. The idea is that task A is classified as subjectively more complex than task B if the probability of correctly solving A is smaller than the probability of correctly solving B for any reward. We provide a full characterization of the incomplete order over the set of tasks that this criterion induces. This characterization implies that task A will be classified as subjectively more complex than task B if and only if A is both more difficult and much less known than B. This insight is consistent with the general idea within economics that complexity has both an objective and a subjective part. It is also aligned with the literature in psychology and information science which —unlike economics— have identified prior uncertainty as a key dimension of complexity. Then, using a lab experiment, where we can exogenously control both difficulty and uncertainty, we corroborate our theoretical predictions. Thus, the recently surging use of expected accuracy in economics as a proxy for subjective complexity is well warranted, as long as expected accuracy is elicited for multiple different rewards using the strategy method.

Keywords: subjective complexity, expected accuracy, pure difficulty, prior uncertainty, reward, rational inattention.
JEL codes: D83, D90.

*An earlier draft of the paper circulated under the title "A robust measure of complexity". We are greatly indebted to Larbi Alaoui, Mark Dean, David Gill, Duarte Goncalves, Paul Grass, Hans-Theo Normann, Antonio Penta, Andy Schotter, Martin Strobel, Peter Wakker, Georg Weizsäcker and the audience in Barcelona (Workshop on Preferences, Bounded Rationality, and Strategic Interactions), Berlin (Behavioral Economics Seminar Series), Belgium Luxembourg Netherlands Theory Seminar (BLuNT), Paris (LEMMA), Rabat (ASSET) and Maastricht (BEELab meeting) for very helpful comments and valuable suggestions.

†Homepage: egorbronnikov.github.io; E-mail: egor.bronnikov@maastrichtuniversity.nl

‡Homepage: www.elias-tsakas.com; E-mail: e.tsakas@maastrichtuniversity.nl

1. Introduction

Task complexity is a fundamental concept across numerous scientific domains, including computer science, psychology, neuroscience, information science, etc.[1] More recently, its importance has also been recognized by (behavioral) economists as a key predictor of mistakes in several choice domains (e.g., see Oprea, 2024a, and references therein). At the same time, complexity has been a rather elusive concept, in the sense that despite systematic attempts to provide a formal definition (e.g., Gabaix and Graeber, 2024; Goncalves, 2024; Oprea, 2024a; Shubbat and Yang, 2024), it is still typically understood in a casual way without consensus on what exactly it means.

As a response to the lack of a universally accepted definition, many economists have taken a different approach, focusing on measures which can serve as proxies for subjective complexity, without explicitly specifying what the latter means (see Oprea, 2024a, for a review of such measures). The one measure that has recently surged in the literature is expected accuracy, i.e., the subjective probability that an agent attaches to the task being solved correctly (Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Agranov *et al.*, 2025).[2] Within this literature, the implicit assumption is that the agent will perceive task A as subjectively more complex than task B, whenever the subjective probability of correctly solving A is larger than the subjective probability of correctly solving B.

The appeal of this approach is that expected accuracy is intuitive, simple and portable across tasks. At the same time, this stream of literature is primarily experimental, with relatively little attention on the theoretical foundations. As a result, it still is not clear what exactly we measure with expected accuracy, and a fortiori whether expected accuracy is indeed a good proxy for subjective complexity.

In this paper, we begin by formalizing expected accuracy within a standard rational inattention framework (e.g., Sims, 2003; Matějka and McKay, 2015; Caplin *et al.*, 2022), which can be embedded in a broader class of models that allow us to study tradeoffs between benefits from thinking through the task and cognitive costs (Alaoui and Penta, 2022). The advantage of this approach is that, although it is reduced-form, it is flexible enough to accommodate most channels that naturally connect expected accuracy with subjective complexity.

The primary such channel is pure difficulty. This is an objective parameter that summarizes all task-specific characteristics that determine how costly it is to think through the task.[3]

[1]Throughout the paper, we focus on cognitive tasks which can take the form of a choice problem with one correct/rational action.

[2]This measure has appeared in the literature in different forms and under different terms, e.g., as cognitive uncertainty (Enke and Graeber, 2023) or confidence (Agranov *et al.*, 2025).

[3]Within economics, pure difficulty is often treated as a synonym for complexity (e.g., Goncalves, 2024). In order to reconcile this earlier work with our terminology, we will also refer to pure difficulty as objective complexity, which is narrower than the notion of subjective complexity. This distinction has also been made by Agranov *et al.* (2025).

Formally, it takes the form of a multiplier in the cost of information (e.g., Dean and Neligh, 2023; Goncalves, 2024). It is a channel linking expected accuracy to subjective complexity in the sense that, ceteris paribus, more difficult tasks are both subjectively more complex and less likely to be solved.

The second major channel is prior uncertainty. This is a subjective task-specific parameter that summarizes prior knowledge, experience, familiarity for the task. It is formally captured by the Shannon entropy of the prior belief about the correct answer (e.g., Cover and Thomas, 2006). It constitutes a channel linking expected accuracy with subjective complexity in the sense that, ceteris paribus, tasks that are a priori better known are both subjectively less complex and more likely to be solved. While prior uncertainty is new in the economics literature on complexity, in other disciplines it has been already identified as one of the main dimensions of complexity, e.g., see review articles from psychology (Campbell, 1988) and information science (Vakkari, 2005).

But, at the same time, expected accuracy also depends on parameters that are not related to subjective complexity. These are exogenous parameters that affect the effort/attention that the agent puts in solving the problem, without being correlated with any reasonable notion of subjective complexity. The main such parameter is the reward for solving a task correctly, viz., a higher reward makes it more likely that a task is solved correctly, while it has nothing to do with how complex the task is deemed.[4]

As a result, we cannot be sure whether differences in expected accuracy reflect differences in parameters that relate to subjective complexity (viz., pure difficulty and prior uncertainty), or whether they are driven by unrelated parameters (viz., our choice of the reward). In fact, our model predicts that the latter will often be the case, i.e., sometimes, task A will be more likely to be solved than task B when the reward is small, and vice versa when the reward is large. Obviously, in such cases, expected accuracy is a bad proxy for ranking A and B in terms of subjective complexity.

Of course this does not mean that expected accuracy is a bad proxy for subjective complexity altogether, but rather that it is a bad proxy for specific pairs of tasks that exhibit a reversal of the type we just described. That is, we are comfortable in proxying subjective complexity with a uniform version of expected accuracy:

> *Task A will be classified as subjectively more complex than task B whenever, for every reward the subjective probability of correctly solving A is smaller than the subjective probability of correctly solving B.*

Obviously, this is a strong dominance criterion, meaning that the resulting (uniform expected accuracy) order will be incomplete. But the question is "how much incomplete" it is.

Our first main result provides a full characterization of our uniform expected accuracy order in terms of the two task-dependent parameters in our model, viz., pure difficulty and

[4]Other such parameters are discussed in Section 5.1.

prior uncertainty (Theorem 1). In particular, our result implies that task A is classified as subjectively more complex than task B if and only if A is both more difficult and much less known than B. In this sense, our characterization is analogous to vector utility representations of incomplete preferences (Ok, 2002).

Note that conceptually our main point is not to merely conclude that expected accuracy is linked with pure difficulty and prior uncertainty.[5] Rather, our key insight is to go a step further and establish that expected accuracy can be used as a proxy for subjective complexity, when we look at the subset of tasks where pure difficulty and prior uncertainty are comonotonic. In this sense, our result reinforces the idea that reasonable notions of subjective complexity have simultaneously both objective and subjective dimensions (Oprea, 2024a).

Now, let us focus on pairs of tasks that are not ranked by our uniform expected accuracy order, i.e., task A is more difficult than B, but B is much less known than A. Then, according to our previous result, sometimes A is deemed less likely to be solved (viz., pure difficulty is the dominant force), and other times B is deemed less likely to be solved (viz., prior uncertainty is the dominant force). As it turns out, there is a single-crossing condition that describes which is the dominant force for each reward. In particular, for small rewards the difficult and better known task is deemed more likely to be solved (viz., task A), whereas for large rewards the easy and less known task is deemed more likely to be solved (viz., task B), where small and large rewards are separated by a single threshold (Theorem 2).

The intuition is quite clear. For small rewards it is not worth it to put much cognitive effort in solving the task, and therefore one tends to rely mainly on prior knowledge. On the other hand, for large rewards it pays off to think through the problem carefully, and therefore one will rely more on cognitive effort, which in turn is more effective when the task is easy. The latter illustrates the mechanism through which rewards moderate the relationship between expected accuracy and subjective complexity.

Now that we have established when expected accuracy is a good proxy for subjective complexity, we observe that it can also be used to identify differences in pure difficulty. This is because there are two possible scenarios: either A is deemed more likely to be solved than B *for every reward*, in which case B will be more difficult (by Theorem 1), or A is deemed more likely to be solved than B *only for small rewards*, in which case A will be more difficult (by Theorem 2).[6]

In the second part of the paper, we test the predictions of our theoretical model in a lab experiment, where we can exogenously control and manipulate the two dimensions of subjective complexity that we previously identified, viz., pure difficulty and ex ante uncertainty.

Our experiment will consist of two stages. In the first stage —which is merely auxiliary for

[5]It has been already pointed out in the literature that accuracy is linked with pure difficulty (Goncalves, 2024) and with prior uncertainty (Fudenberg *et al.*, 2018).

[6]A different approach for eliciting objective complexity by manipulating the underlying reward is used in Goncalves (2024) and Goncalves *et al.* (2024).

the second stage— the setting is similar to the one in Dean and Neligh (2023), i.e., subjects see a screen with blue and red balls scattered across, and they are asked to guess the dominant color.[7] Pure difficulty is determined by the total number of balls on the screen, and prior uncertainty is determined by the probability of the screen containing more blue balls, where this probability is given to them by the experimenter. A combination of difficulty and uncertainty characterizes a task. We run three treatments (viz., no-stakes, low-stakes, high-stakes) between subjects, which only differ in the bonus the subjects receive for guessing correctly. Then, in the second —and main— stage of the experiment we ask another group of subjects from the same pool to guess, for each task and each reward, the proportion of first-stage subjects that answered correctly.[8]

All hypotheses that are derived from our two main theorems are confirmed. Starting with Theorem 1, we find that for any pair of tasks, A and B, for which A (weakly) dominates B both in pure difficulty and prior uncertainty, second-stage participants reported significantly lower probability that the dominant color was guessed correctly in task A than in task B, in each of the treatments. Then, moving to Theorem 2, we consider pairs of tasks, A and B, such that A dominates B in pure difficulty, and B dominates A in prior uncertainty. Then, as predicted by Theorem 2, we find that whenever second-stage participants reported that A is more likely to be solved than B in the high-stakes treatment, they also reported it to be the case in the low-stakes treatment. Likewise, whenever second-stage participants reported that B is more likely to be solved than A in the low-stakes treatment, they also reported it to be the case in the high-stakes treatment. The aforementioned results essentially constitute a validation of uniform expected accuracy as a good proxy for subjective complexity.

Summarizing our key theoretical contribution, we show that expected accuracy is a good proxy for ranking tasks in terms of subjective complexity if pure difficulty and prior uncertainty agree on how the tasks are ordered. This idea is consistent with the general idea within economics that complexity has both an objective and a subjective part (Oprea, 2024a), and is also aligned with the understanding of complexity in other disciplines (Campbell, 1988; Vakkari, 2005).

Then turning to our practical contribution, our work supports the recently surging experimental literature that uses expected accuracy to define subjective complexity (Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Agranov *et al.*, 2025). And while we conclude that more detailed data are needed (using the strategy method across the reward space), we also show that the initial intuition —of using expected accuracy— was in the right direction.

The literature on complexity is vast, and as such we are de facto forced to make a selection

[7] In recent work, Goncalves *et al.* (2024) also employ a similar design to study the sensitivity of choices to incentive changes. We further elaborate on their work later on.

[8] Note that we use as a proxy for their own expected accuracy, their beliefs about the expected accuracy of other individuals similar to them. The reason is to avoid potential hedging behavior (Blanco *et al.*, 2010). This idea is similar to the one employed in the literature on Bayesian markets (Baillon, 2017). We further elaborate on practical issues that pertain to elicitation in Section 5.4.

of what in our view is the most relevant subset. We will not even attempt to touch the related literatures within other disciplines, such as computer science or cognitive sciences.

Early work focused primarily on the role of strategy complexity within game theory (e.g., for early contributions, see Rubinstein, 1986; Abreu and Rubinstein, 1988). More recently, the focus has shifted towards understanding the role of complexity in well-established behavioral patterns, e.g., how does complexity interact with risk attitudes (Oprea, 2024b), ambiguity attitudes (Aydogan *et al.*, 2023), time preferences (Enke *et al.*, 2024a). As a response to these links, there are several attempts to formalize a definition 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 (Callander, 2011; Fehr and Rangel, 2011; Goncalves, 2024), similarly to what is often done in psychometrics. Alternative approaches use tradeoffs (Shubbat and Yang, 2024) or degrees of contingent reasoning (Nagel and Saitto, 2025) to define complexity. The common denominator throughout these papers is that the respective definitions of complexity are input-based, i.e., complexity is defined by means of some characteristics of the environment.

What is closer to our work is the literature on measures that serve as proxies for complexity without needing to provide an explicit definition. As Oprea (2024a) elegantly points out, this literature can be classified into three large streams, depending on the measurement tool. Within the first stream, we encounter direct measures, such as willingness to pay in order to avoid dealing with a certain task (Oprea, 2020), response times (Gill and Prowse, 2023; 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 most commonly beliefs about one's own accuracy (Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Agranov *et al.*, 2025). This last paper of Agranov *et al.* (2025) is also conceptually close to our work in that they make explicit the distinction between objective and subjective complexity.

The link between accuracy and complexity has been already studied in the literature, albeit from a reverse angle. Namely, instead of using accuracy as a proxy for complexity without being explicit about a definition of complexity, Goncalves (2024) starts by fixing a reasonable notion of objective complexity and proceeds to show that accuracy is decreasing in complexity. He does so in two different models, viz., a drift-diffusion model where complexity is defined as the signal-to-noise ratio, and in a rational-inattention model —similar to ours— where complexity is defined as pure difficulty (in our terminology). Then, using this structural framework, Goncalves *et al.* (2024) conduct an experiment —using a design from which we have also shared many features— where they leverage (small) changes in incentives to infer the exogenously controlled complexity parameter (viz., the signal-to-noise ratio) from choices.

Throughout the economics literature on complexity —explicitly or implicitly— prior uncertainty is exogenously shut down as a potential channel of complexity. And this can potentially explain why the role of the underlying reward has not been identified. For example, in the second part of Goncalves (2024) where a rational inattention model is used (similar to ours), accuracy is decreasing in complexity for any reward. However, as we have already mentioned, in this paper complexity is synonymous to pure difficulty, meaning that prior uncertainty is kept exogenously fixed across tasks. Hence, this is a very special case of our setting. To the best of our knowledge, the only place in the economics literature where the relationship between accuracy and prior uncertainty is studied is the work of Fudenberg *et al.* (2018), which in a drift-diffusion model shows that accuracy is larger when decisions are made quickly subject to prior uncertainty be low. However, this work does not relate accuracy with complexity.

Somewhere in between the input-based and the measure-based approach one finds a stream of literature that uses as parameters for complexity lotteries characteristics (Huck and Weizsäcker, 1999; Fudenberg and Puri, 2022; Enke and Shubatt, 2024; Hua Hu, 2023).

Related to our work is the strand of the literature stemming from the recent work of Avoyan and Schotter (2020) and Avoyan *et al.* (2023). The main idea is that the agent is given a time budget to allocate between two tasks. The distribution of the time budget can be potentially used as an alternative proxy for subjective complexity. Of course, the time allocation will depend on the underlying reward. Then, one can leverage observed budget allocations for different rewards to obtain a good proxy for subjective complexity. This approach would also relate to the work of Goncalves *et al.* (2024), who also leverage exogenous reward manipulations, as we have already discussed above.

This entire literature is somewhere in the intersection of two strands of the surging field of Cognitive Economics (Alaoui and Penta, 2022; Caplin, 2025; Enke, 2024), which also incorporates topics such as rational inattention, cognitive uncertainty, etc. Specifically related to our work, within this broader literature, are the papers that study preference for simplicity (Puri, 2025; de Clippel *et al.*, 2025; Mononen, 2025, and references therein) and model uncertainty (Mussolf and Zimmermann, 2025, and references therein).

The paper is structured as follows: In Section 2 we introduce our theoretical framework. In Section 3 we introduce our measure of complexity and prove our main characterization results. In Section 4 we present our experiment. In Section 5 we discuss potential extensions and limitations of our theoretical model. Section 6 concludes. Proofs and additional results are relegated to the Appendix.

2. Guessing tasks

2.1. Task characteristics

Consider a binary state space $S = \{s_0, s_1\}$ and a (female) agent. The associated task requires the agent to guess the state realization.

Definition 1. Formally, a task is identified by the state space S itself. The set of all possible tasks is denoted by $\mathcal{S}_0$. ◁

The agent is assumed to be a SEU maximizer who weakly prefers guessing correctly, and strictly prefers more money over less money. In particular, let $X := [0, \infty)$ be the set of monetary rewards that the agent can potentially receive for guessing correctly. Let $v(x) \geq 0$ denote her (Bernoulli) utility from receiving a reward $x \in X$ upon answering correctly. We assume that v is continuously strictly increasing and unbounded. Her utility from guessing wrongly is normalized to 0.[9]

For an arbitrary task $S \in \mathcal{S}_0$ and a reward $x \in X$, the agent's guess will be denoted by $r^x \in S$. Formally, it is an act that yields utility $v(x)$ if she guesses the true state and 0 otherwise. So, for an arbitrary belief that attaches subjective probability $q \in (0,1)$ to s_1, her expected utility is given by

$$\mathbb{E}_q(v(r^x)) = \begin{cases} (1-q)v(x) & \text{if } r^x = s_0, \\ qv(x) & \text{if } r^x = s_1. \end{cases}$$

Thus, her indirect expected utility is given by

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

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

The agent's prior belief is identified by the subjective probability $\mu_S \in (0,1)$ that she assigns to s_1. We measure her degree of *prior uncertainty* about S by the Shannon entropy

$$H(\mu_S) = -\mu_S \log \mu_S - (1-\mu_S) \log(1-\mu_S), \tag{2}$$

of the prior belief. This parameter summarizes the agent's prior information, knowledge, experience, etc. It takes values in $(0,1]$, and it is strictly decreasing with respect to the distance from the uniform belief. This notion of prior uncertainty has solid foundations within information theory (Cover and Thomas, 2006). Importantly, for the purposes of our paper, the only thing that matters is the ordinal relation that it induces over the set of priors, and thus it is without loss of generality to take any strictly increasing transformation of $H(\mu_S)$.

[9] Later we generalize our framework to allow for task-dependent utilities (Section 5.2) and state-dependent utilities (Section 5.3).

Before making a guess, the agent decides how much attention to allocate in solving the task (e.g., Sims, 2003; Caplin *et al.*, 2022). An attention strategy is described in its reduced form by a Bayesian signal, which is uniquely identified by a mean-preserving distribution of posterior probabilities (Kamenica and Gentzkow, 2011). The set of all signals for task S is denoted by

$$\Pi_S = \Big\{ \pi \in \Delta\big([0,1]\big) : \mathbb{E}_\pi(q) = \mu_S \Big\} .$$

For any signal $\pi \in \Pi_S$, define the agent's ex ante indirect expected utility,

$$G^x(\pi) := \mathbb{E}_\pi\big(g^x(q)\big). \tag{3}$$

As usual, we assume that attention is costly. The cost function is assumed to be uniformly posterior separable (Caplin *et al.*, 2022; Tsakas, 2020), i.e., there is a strictly concave function $c : [0,1] \to \mathbb{R}$ such that the cost of signal $\pi \in \Pi_S$ is given by

$$C_S(\pi) = \kappa_S\Big(c(\mu_S) - \mathbb{E}_\pi\big(c(q)\big)\Big), \tag{4}$$

where the function c represents the agent's ability to solve tasks, and $\kappa_S > 0$ is a parameter of the task's *pure difficulty*. Note that the cost consists an objective task characteristic (viz., the parameter κ_S) and a subjective individual characteristic (viz., the function c). Posterior-separable costs functions have solid foundations (Denti, 2022; De Oliveira and Mensch, 2025) and are supported by experimental evidence (Dean and Neligh, 2023). Without loss of generality, c is assumed to be twice continuously differentiable in $(0,1)$. 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 there exists some linear function $t(q)$ such that, for every $q \in [0,1]$, we have $c(q) + t(q) = c(1-q) + t(1-q)$. For an axiomatization of symmetric cost functions, see Hébert and Woodford (2021). We further discuss the symmetry assumption in Section 5.5.

Remark 1. Overall, in our model, there are two task-specific parameters: pure difficulty and prior uncertainty. Pure difficulty is objective (i.e., it can be assumed to be common for all agents), whereas prior uncertainty is subjective (i.e., it typically varies across agents). Crucially, the reward x is not a task characteristic, but rather a directly observable parameter which is controlled by the analyst, and can vary arbitrarily within the same task S. $\triangleleft$

2.2. Optimal attention

Throughout this section, fix a task $S \in \mathcal{S}_0$. For each given reward $x \in X$, the agent faces a tradeoff, in that more informative signals help her achieve higher expected utility, but at the same time are also more costly. Formally, the agent solves the following optimization problem:

$$\max_{\pi\in\Pi_S}\Big(G^x(\pi) - C_S(\pi)\Big). \tag{5}$$

It is not difficult to verify that there is a unique optimal signal, henceforth denoted by π_S^x, which is found by a standard concavification argument (e.g., Kamenica and Gentzkow, 2011; Matějka and McKay, 2015). This signal is characterized by the threshold q_S^x, which is the unique solution in $[0, 1/2]$ to

$$c'(q) = \min\left\{\frac{v(x)}{\kappa_S}, c'(0)\right\}. \tag{6}$$

In particular, if $\mu_S \leq q_S^x$ or $\mu_S \geq 1-q_S^x$, the optimal signal π_S^x will be completely uninformative, putting probability 1 to the prior μ_S itself. On the other hand, if $q_S^x < \mu_S < 1-q_S^x$, the optimal signal π_S^x will distribute all its probability mass between q_S^x and $1-q_S^x$. Throughout the paper, we will denote the Shannon entropy of q_S^x by

$$H(q_S^x) = -q_S^x \log q_S^x - (1-q_S^x)\log(1-q_S^x). \tag{7}$$

It is not difficult to verify that π_S^x is completely uninformative if and only if $H(\mu_S) \leq H(q_S^x)$. That is, $H(q_S^x)$ is the uncertainty threshold for acquiring information through attention when the reward is x. Intuitively, for large prior uncertainty, it makes sense to pay attention, thus reducing uncertainty to $H(q_S^x)$. On the other hand, for small prior uncertainty, it is not worth to pay any attention, thus remaining with the prior uncertainty.

Of particular importance for our results will be $H(q_S^0)$, i.e., the uncertainty threshold in the absence of any reward. Note that, ceteris paribus, $H(q_S^0)$ is continuously increasing with respect to pure difficulty, i.e., $\kappa_S \geq \kappa_{S'}$ implies $H(q_S^0) \geq H(q_{S'}^0)$. Furthermore, it is the case that $H(q_S^0) \to 0$ as $\kappa_S \to 0$, and $H(q_S^0) \to 1$ as κ_S grows arbitrarily large. Finally, if the agent cares only about money and does not have any intrinsic incentives, we will trivially obtain $H(q_S^0) = 1$ for every $s \in \mathcal{S}_0$.

3. A measure of subjective complexity

3.1. Uniform expected accuracy

Within our model of rational inattention we can formalize the notion of expected accuracy, which has been extensively —albeit informally— used across the experimental literature as a proxy for subjective complexity (e.g., Agranov *et al.*, 2025; Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Oprea, 2024a).

The main advantage of our model is that it is flexible enough to encompass different channels that naturally connect expected accuracy with subjective complexity. In particular, it accommodates not only pure difficulty, but also prior uncertainty. In disciplines outside Economics, the latter has been recently recognized as a key dimension of subjective complexity (e.g., for standard review articles, see Campbell, 1988; Vakkari, 2005). The idea is quite clear: tasks which are better known a priori are perceived both more likely to be solved correctly and less complex.

Formally, *(subjective) expected accuracy given the reward* x is equal to the total probability of guessing correctly. This is given by the formula

$$P(S,x) = \max\left\{\mu_S, 1-\mu_S, 1-q_S^x\right\}, \tag{8}$$

and it is illustrated in Figure 1.[10] Intuitively, whenever prior uncertainty is high (i.e., $H(\mu_S) > H(q_S^x)$), the agent will acquire information using the attention strategy π_S^x, which in turn will induce one of the two posterior beliefs in $\text{supp}(\pi_S^x) = \{q_S^x, 1-q_S^x\}$. Regardless which of the two posteriors is eventually realized, the probability of the optimal guess being correct is equal to $1-q_S^x$, and therefore her overall subjective accuracy will be equal to $P(S,x) = 1-q_S^x$. On the other hand, whenever prior uncertainty is low (i.e., $H(\mu_S) \leq H(q_S^x)$), the agent will not pay any attention, and she will pick her optimal guess given the prior belief. Thus, her subjective accuracy will be equal to $P(S,x) = \max\{\mu_S, 1-\mu_S\}$.

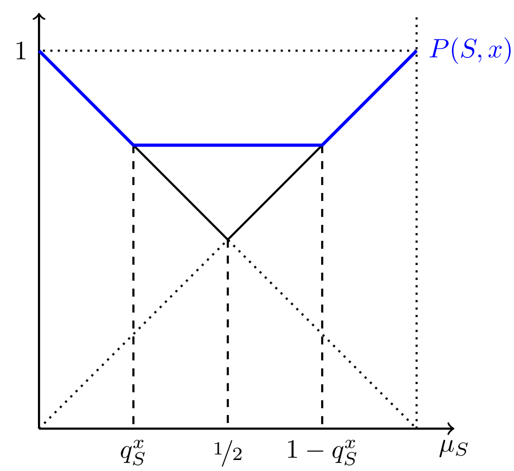


Figure 1: The blue piecewise linear function is the expected accuracy (as a function of the agent's prior), assuming that the optimal signal π_S^x has been used.

Expected accuracy has several compelling features. First of all, it is simple and portable across contexts. Moreover, aligned with our intuition, it is ceteris paribus decreasing in pure difficulty as well as decreasing in prior uncertainty. However, there is caveat: the expected accuracy order often depends on the underlying reward, i.e., there are tasks $S, S' \in \mathcal{S}_0$ such that $P(S,x) > P(S',x)$ and $P(S,x') < P(S',x')$ for separate $x, x' \in X$. In such cases, it is unclear which reward we should use to rank tasks, and therefore it is unclear which of the two tasks should be classified as subjectively more complex. Let us illustrate such a case with an example.

[10] Note that $P(S,x)$ is the probability that the agent attaches ex ante to correctly solving the task, i.e., before she has actually implemented her (optimal) attention strategy. On the other hand, in the experimental literature, where this idea has been put in practice, subjective accuracy is measured ex post, i.e., after a guess has been made. As we discuss in Section 5.6, this distinction does not make any theoretical difference, as long as we fix a single reward x.

Example 1. Consider an agent whose utility for answering correctly is given by $v(x) = 1 + x$, while her cost for acquiring information is given by $c(q) = q - q^2$. Then, for an arbitrary task $S \in \mathcal{S}_0$ with pure difficulty parameter $\kappa_S > 0$ and reward $x \geq 0$, we obtain

$$q_S^x = \max\left\{\frac{1}{2} - \frac{1+x}{2\kappa_S}, 0\right\}.$$

Then, take two tasks $S, S' \in \mathcal{S}_0$, such that S is difficult and relatively known (viz., $\kappa_S = 50$ and $\mu_S = 65\%$), while S' is easy and completely unknown (viz., $\kappa_{S'} = 5$ and $\mu_{S'} = 50\%$). Then, for a small reward (e.g., $x = 0$) we have $q_S^0 = 49\%$ and $q_{S'}^0 = 40\%$, whereas for a larger reward (e.g., $x = 1$) we have $q_S^1 = 48\%$ and $q_{S'}^1 = 30\%$. Hence, by (8), we obtain

$$\begin{aligned} P(S,0) = 65\% \quad &> \quad 60\% = P(S',0), \\ P(S,1) = 65\% \quad &< \quad 70\% = P(S',1), \end{aligned}$$

meaning that her expected accuracy order depends on the underlying reward. ◁

The reason behind this type of ambiguity can be traced back to the fact that we allow prior uncertainty to differ across tasks, meaning that there are two task specific parameters that affect expected accuracy, viz., pure difficulty and prior uncertainty. Then, the weight of each parameter in determining expected accuracy varies in the reward, e.g., in Example 1, for $x = 0$ prior uncertainty is the decisive parameter, whereas for $x = 1$ it is pure complexity that decides which task is more likely to be solved. Note that this issue will be further exacerbated if we add more individual parameters that affect expected accuracy, but are not related to subjective complexity (see Section 6).

This type of ambiguity does not cancel out all the appealing features of expected accuracy, but it forces us to be more conservative when we use it a proxy for subjective complexity. In practice, we introduce the uniform expected accuracy order, according to which two tasks can be compared if one is more likely to be solved than the other for any possible reward.

Definition 2. (Uniform Expected Accuracy order). For any two tasks $S, S' \in \mathcal{S}_0$, we write $S \succeq S'$ whenever

$$P(S,x) \leq P(S',x) \tag{9}$$

for all $x \in X$. In this case, we classify task S as *subjectively more complex* than task S'. The asymmetric and the symmetric parts of $\succeq$ are defined as usual. ◁

Note that $\succeq$ will be incomplete, as it is induced by a strong dominance criterion. This is the price that we have to pay for being comfortable in using expected accuracy as a proxy for subjective complexity, without worrying about the aforementioned ambiguity problem. Then, our main question focuses on characterizing the pairs of tasks which are $\succeq$-comparable. That is, we want to know the extent to which expected accuracy is an unambiguously good proxy of subjective complexity.

3.2. Characterization results

A task $S \in \mathcal{S}_0$ is said to be trivial if the optimal signal π_S^x reveals the true state with certainty (i.e., $q_S^x = 0$) for every $x \geq 0$. Obviously, if a task S is trivial, then it is subjectively simpler than any other task, as it will satisfy $P(S, x) = 1$ for every $x \in X$. The set of non-trivial tasks is henceforth denoted by $\mathcal{S} \subseteq \mathcal{S}_0$, and it is characterized by a difficulty threshold.

Proposition 1. *There is some $\kappa \geq 0$ such that,*

$$S \in \mathcal{S} \;\Leftrightarrow\; \kappa_S > \kappa. \tag{10}$$

The idea is quite simple: in order for a task to be non-trivial, the attention 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 this sense, if the agent cares only about monetary payoffs and does not get any intrinsic utility from solving the task correctly, the threshold becomes $\kappa = 0$.

In our main result below, we characterize how non-trivial tasks are ranked if we use our uniform expected accuracy measure. Our characterization of $\succeq$ within $\mathcal{S}$ has the same structure as the vector-valued utility representation of incomplete preference relations in Ok (2002). In particular, let us first define the vector-valued function $\boldsymbol{\phi} : \mathcal{S} \to \mathbb{R}^2$, where

$$\boldsymbol{\phi}_1(S) := \kappa_S \text{ and } \boldsymbol{\phi}_2(S) := \min\left\{H(\mu_S), H(q_S^0)\right\}. \tag{11}$$

Let $\geq$ be the usual order over $\mathbb{R}^2$, i.e., we will write $\boldsymbol{\phi}(S) \geq \boldsymbol{\phi}(S')$ if and only if $\boldsymbol{\phi}_1(S) \geq \boldsymbol{\phi}_1(S')$ and $\boldsymbol{\phi}_2(S) \geq \boldsymbol{\phi}_2(S')$. Then, we are ready to obtain our main characterization result.

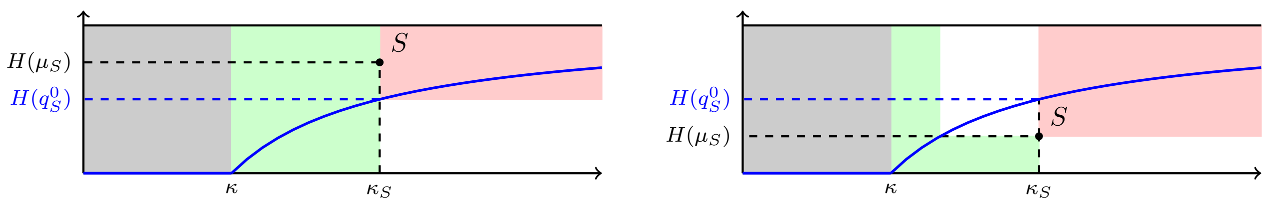
Theorem 1 (Main characterization result)**.** *For any pair $S, S' \in \mathcal{S}$:*

$$S \succeq S' \;\Leftrightarrow\; \boldsymbol{\phi}(S) \geq \boldsymbol{\phi}(S'). \tag{12}$$

Graphically, the previous result is illustrated in Figure 2 below. The two subfigures correspond to the two values that $\boldsymbol{\phi}_2(S)$ can potentially take. On the lefthand side prior uncertainty is high, i.e., we have $H(\mu_S) > H(q_S^0)$, which in turn means that $\boldsymbol{\phi}_2(S) = H(q_S^0)$. On the righthand side prior uncertainty is low, i.e., we have $H(\mu_S) \leq H(q_S^0)$, and therefore we obtain $\boldsymbol{\phi}_2(S) = H(\mu_S)$.

Remark 2. If the agent cares only about monetary payoffs without having any intrinsic incentives, we obtain $H(q_S^0) = 1$ for every $S \in \mathcal{S}_0$, meaning that no task is trivial, and a fortiori only part (b) of Figure 2 is relevant. Hence, we have $\boldsymbol{\phi}_2(S) = H(\mu_S)$ for every $S \in \mathcal{S}$. ◁

The intuition of the previous result is quite obvious: in order to classify a task to be subjectively more complex (than another task) using expected accuracy, it is not sufficient that it is more difficult; it also needs to be sufficiently less known a priori. The latter highlights the



(a) High prior uncertainty: $H(\mu_S) > H(q_S^0)$.

(b) Low prior uncertainty: $H(\mu_S) \leq H(q_S^0)$.

Figure 2: In both cases, the red area contains the tasks that are subjectively more complex than S, and the green area are those that are subjectively simpler than S. Finally, the grey area contains the trivial tasks, meaning that they are subjectively simpler than every other task, including S. Finally, the white area contains the tasks that are not $\succeq$-comparable to S.

importance of prior uncertainty as a new dimension of complexity, consistently with existing literature in other disciplines, such as psychology (Campbell, 1988) and information science (Vakkari, 2005). Moreover, it reinforces the idea that subjective complexity consists of both objective and a subjective dimensions (Oprea, 2024a).

The fact that $\succeq$ is incomplete also distinguishes our notion of subjective complexity from most of the existing literature, which typically considers complete complexity orders, regardless whether complexity is explicitly defined (e.g., Goncalves, 2024) or proxied by some measure (Oprea, 2024a, and references therein). In fact, complete complexity orders are used even when subjective complexity is proxied by more than one dimensions (Agranov *et al.*, 2025).[11]

The way in which $\succeq$ is incomplete in our case is not arbitrary, i.e., every pair of tasks that are not $\succeq$-comparable is characterized by two thresholds (in the reward) space which determine the task in which the agent is expected to be more accurate for different rewards.

Theorem 2 (Single-crossing)**.** *Suppose that $S, S' \in \mathcal{S}$ are not $\succeq$-comparable, in that $\boldsymbol{\phi}_1(S) > \boldsymbol{\phi}_1(S')$ and $\boldsymbol{\phi}_2(S) < \boldsymbol{\phi}_2(S')$. Then, there exist two thresholds $0 < x_1 < x_2 \leq \infty$ such that:*

(i) $P(S, x) > P(S', x)$ *for all* $x \in (0, x_1)$,

(ii) $P(S, x) < P(S', x)$ *for all* $x \in (x_1, x_2)$,

(iii) $P(S, x) = P(S', x)$ *for all* $x \in (x_2, \infty)$.

The intuition is quite simple. For small rewards, it is not worth paying attention to the tasks, and therefore the agent will mainly rely on her prior knowledge/experience/information. As a result, it is more likely that she guesses correctly in the task which is a priori better known, even though it is more difficult. On the other hand, for large rewards, it pays off to think through the tasks, and therefore she will primarily rely on information she acquires through attention. Thus, she is more likely to guess correctly in the easy task, despite the

[11] In their paper, Agranov *et al.* (2025) define a subjective complexity order at the individual level by means of a discrete measure of subjective accuracy (given a single reward) and a measure of effort.

fact that a priori much less is known about it. And of course, for very large rewards, she will certainly guess correctly in both of them.[12] The previous theorem is graphically illustrated in Figure 3 below.

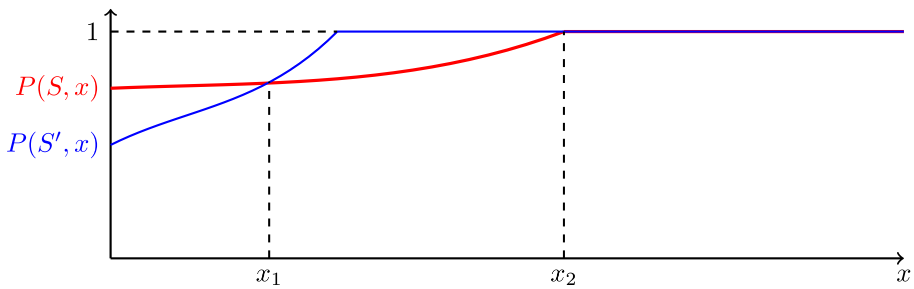


Figure 3: The blue task S' is easier and a priori more uncertain than the red task S. This is why, for small rewards (viz., $x < x_1$) the agent relies more on her prior knowledge/experience/information and is therefore more likely to solve S. On the other hand, for large rewards (viz., $x_1 < x < x_2$) she relies more on her attention, and hence she is more likely to solve S'. Finally, for very large rewards (viz., $x \geq x_2$) she will solve both tasks with certainty.

While our work is about providing a proxy for subjective complexity, it also allows us to rank tasks with respect to their objective complexity, viz., pure difficulty. In particular, for any two tasks $S, S' \in \mathcal{S}_0$, our theorems allow us to distinguish between the following two cases: for any two tasks, either there is no reversal (Theorem 1), or there is exactly one reversal (Theorem 2) in the expected accuracy order across X. In both cases, we can definitively say which is the more difficult task. Specifically, if $P(S, x) \geq P(S', x)$ for all $x \in X$, then $\kappa_S \leq \kappa_{S'}$. On the other hand, if $P(S, x) > P(S', x)$ for some $x \geq 0$, and $P(S, x') < P(S', x')$ for some $x' > x$, then $\kappa_S > \kappa_{S'}$.

Ours is not the only paper to leverage rewards in order to elicit objective complexity. In particular, Goncalves (2024) uses a different identification strategy which relies on providing additional incentives for one of the two guesses in each task, and then comparing the rate at which the probability of choosing this guess increases. His approach is experimentally validated by Goncalves *et al.* (2024) in a broad range of domains. A key difference to our work is that in their paper they measure the probability of each guess being made, whereas we measure the probability of the selected guess being correct. So, in practice, they do not require the actual state to be ex post observable. But on the flip side, they need to observe stochastic choice data of some sort, whereas we require data obtained with the strategy method for a single realization of each task.

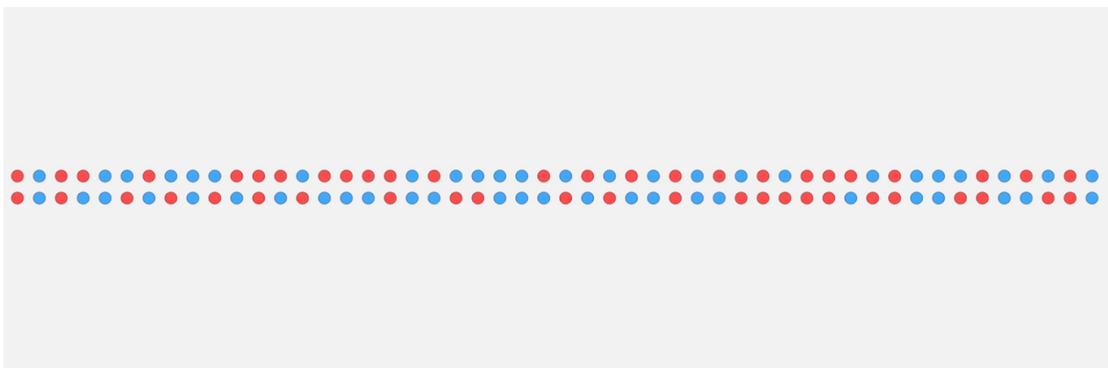
[12]The second threshold x_2 is finite if and only if c does not becomes infinitely steep at the boundaries of $[0, 1]$. The most common example where x_2 is infinite is the case with entropic costs. In such cases, the interval (x_2, ∞) is trivially empty.

4. Proof of concept

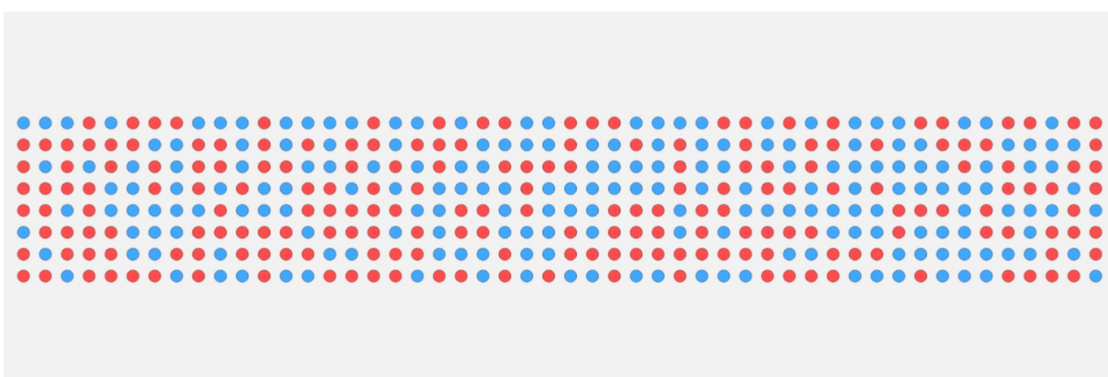
Our theory characterizes the uniform expected accuracy order in terms of the two task-specific parameters (viz., pure difficulty and prior uncertainty), and then uses it as a proxy for subjective complexity. In this section, we are going to test the hypotheses that we derive from our theory regarding $\succeq$ in a lab setting where we can observe and exogenously manipulate the two parameters.

4.1. Experimental design

The experiment is divided in two stages. Both stages were run in the Behavioral and Experimental Economics Lab (BEELab) in Maastricht University, with subjects recruited from the same pool of students.



(a) Easy blue-dominant panel: it contains 49 red and 51 blue marbles.



(b) Difficult red-dominant panel: it contains 201 red and 199 blue marbles.

Figure 4: Examples of two panels that are seen by participants in the first stage.

In the first stage each participant would face a sequence of tasks, which we describe below in detail, and are similar to the ones in Dean and Neligh (2023) and Goncalves *et al.* (2024). One of the tasks would be randomly drawn in the end, and if the participant answered correctly in this task, this participant would receive a bonus payment (€x), on top of the show-up fee (€5). There were two treatments in the first stage which would only differ in the size of the bonus, i.e., the low-stakes treatment where the bonus was $x = 0.50$, and the high-stakes treatment where the bonus was $x = 10$. That is, the tasks that participants would face in each of the two treatments were identical.

The building blocks of a task are panels with blue and red marbles randomly scattered on the screen. There are four types of panels, differing in the color that the majority of the marbles has (viz., the state) and the total number of marbles on the panel (viz., the difficulty). In particular, a panel contains either more red marbles (i.e., it is a red panel) or more blue marbles (i.e., it is a blue panel). Moreover, a panel is either easy (viz., it contains 100 marbles in total: 51 of one color and 49 of the other color) or (viz., it contains 400 marbles in total: 201 of one color and 199 of the other color). Examples of an easy and a difficult panel are shown in Figure 4.

At the beginning of each task, there is a pool of 10 panels. The participant knows whether the task is easy (i.e., all 10 panels in the pool are easy) or difficult (i.e., all 10 panels in the pool are difficult). Moreover, the participant knows the proportion of red/blue panels in the pool. In particular, in each task there will be either high degree of uncertainty (i.e., 5 red and 5 blue panels) or low degree of uncertainty (i.e., 8 panels of one color and 2 panels of the other color). Then, in each task, a panel is randomly chosen from the respective pool and the participant is asked to guess the state, i.e., whether there are more red or blue marbles in this panel. Thus, there are 4 tasks in total, identified by points in Figure 5.

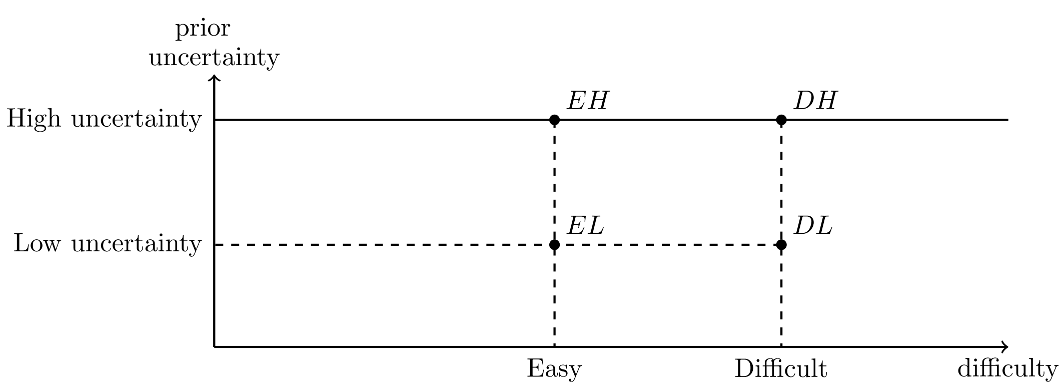


Figure 5: The tasks that the participants face in the first stage of the experiment.

The actual accuracy of participants in the first stage is irrelevant for our hypotheses testing. It will only be used to avoid deception in the second stage, which is the main part of the experiment. As we also discuss in Section 5.4, instead of eliciting participants' beliefs about their own accuracy, we will elicit others' beliefs about the accuracy of participants in the aforementioned first stage. This is exactly what we do in the second stage of our experiment.

The second stage begins by explaining to participants the first stage of the experiment, while explicitly telling them that the first stage was recently conducted with participants drawn from the same pool as themselves. Then, for each task $S \in \{EL, EH, DL, DH\}$ and each reward $x \in \{0.5, 10\}$, they were asked to estimate the proportion of first-stage participants that guessed the correct color.

Formally speaking, what we elicit is the mean of the subjective distribution over the percentages of correct answers $\{0\%, 1\%, \ldots, 100\%\}$, which is henceforth denoted by $P(S, x)$. To guarantee incentive compatibility, we use a standard binarized quadratic scoring rule for eliciting distribution means (Schlag and van der Weele, 2013). The prize of this scoring rule is €8. Notice that, being aligned with Danz *et al.* (2022), the belief elicitation mechanism is shown to the participants only upon request (by clicking on a special button), and otherwise they are simply told that it is in their interest to truthfully report their best estimate.

4.2. Hypotheses and results

Recall that each second-stage participant reports 8 different percentages $P(S, x)$: one for each task $S \in \{EL, EH, DL, DH\}$ and each reward $x \in \{0.5, 10\}$. Our first hypothesis states that participants believe that accuracy is increasing in the incentives.

Hypothesis 1. *For each $S \in \{EL, EH, DL, DH\}$ we hypothesize the following:*

$$P(S, 10) \geq P(S, 0.5). \tag{H.1}$$

This is a sanity check, consistent with the findings in Dean and Neligh (2023).

Our second hypothesis follows directly from Theorem 1, according to which tasks that are simultaneously more difficult and involve more uncertainty are expected to always be less likely to solve, and a fortiori more complex.

Hypothesis 2. *For each $x \in \{0.5, 10\}$ we hypothesize the following:*

$$P(EL, x) \geq P(EH, x) \geq P(DH, x), \tag{H.2a}$$

$$P(EL, x) \geq P(DL, x) \geq P(DH, x). \tag{H.2b}$$

Based on our theory the previous hypothesis should hold irrespective of the attention constraint $\bar{\eta}_S$, in the sense that all tasks that are located north east of S are more complex than S according to our robust measure (Figure 2). The previous hypotheses are also consistent with the findings of Dean and Neligh (2023).

Our third hypothesis focuses on the comparison of two tasks, EH and DL, where DL dominates in one dimension (viz., difficulty) and EH dominates in the other dimension (viz., uncertainty). As a result, it is not obvious whether they are even comparable in the first place: this would depend on the attention constraint $H(q^0_{EH})$, which is of course unobservable. Nonetheless, relying on the single-crossing condition of Theorem 2, we can formulate the following hypothesis.

Hypothesis 3. *We hypothesize the following:*

$$P(EH, 0.5) \geq P(DL, 0.5) \;\Rightarrow\; P(EH, 10) \geq P(DL, 10), \qquad (H.3a)$$
$$P(EH, 10) \leq P(DL, 10) \;\Rightarrow\; P(EH, 0.5) \leq P(DL, 0.5). \qquad (H.3b)$$

Of course, if EH and DL are $\succeq$-comparable, then the hypothesis follows directly from Theorem 1. If on the other hand, they are not $\succeq$-comparable, then in the context of Figure 3 we can set $S = DL$ and $S' = EH$. In this case, $P(EH, 0.5) \geq P(DL, 0.5)$ implies that $x_1 \leq 0.5$, and therefore it must also be the case that $x_1 < 10$. Hence, according to Theorem 2, we should also obtain $P(EH, 10) \geq P(DL, 10)$. The argument is identical for the second part of the hypothesis.

In total, we recruited 56 participants. Their average reported guesses (for the different tasks and rewards) are summarized in Figure 6.

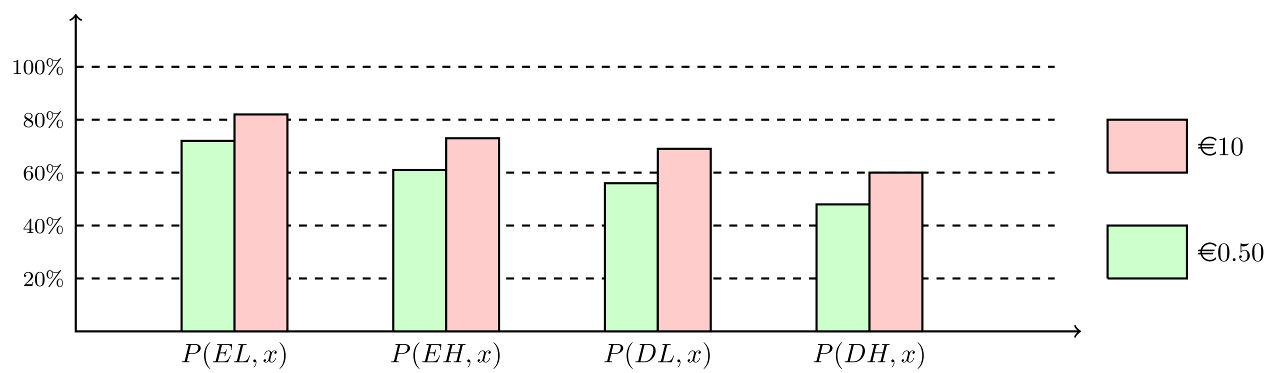
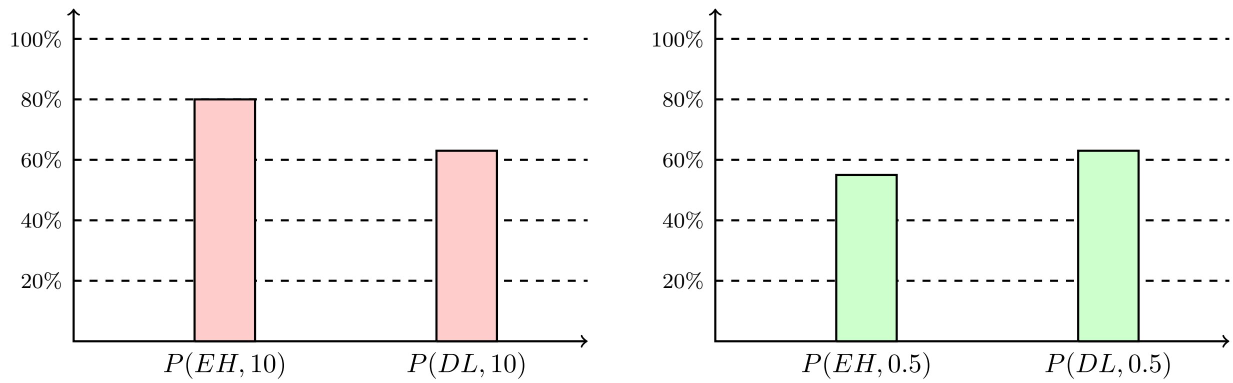


Figure 6: Summary of the expected accuracy reported in the second stage about the performance of participants in the first stage for each task and each reward.

The first observation is that both in Hypothesis 1 and Hypothesis 2, the direction is the one we hypothesized, i.e., higher reward leads on average to higher expected accuracy. Then, using the Wilcoxon signed-rank test, we find that these differences are statistically significant (with $p < 0.01$). This implies that both hypotheses are corroborated.

Then, we turn to Hypothesis 3. To test it, we need to restrict attention to those participants who reported $P(EH, 0.5) \geq P(DL, 0.5)$ for $H.3a$, and respectively to those who reported $P(EH, 10) \leq P(DL, 10)$ for $H.3b$. In each of the two groups, the average accuracies that we respectively observed are the ones depicted depicted in Figure 7.

Obviously, the differences are in the direction that we hypothesized. Then, once again, using Wilcoxon signed-rank test, we find that these differences are statistically significant both in $H.3a$ (with $p < 0.01$) and in $H.3b$ (with $p < 0.03$). Hence, our last hypothesis is corroborated too.



(a) Given $P(EH, 0.5) \geq P(DL, 0.5)$

(b) Given $P(EH, 10) \leq P(DL, 10)$.

Figure 7: Expected accuracy reported in the second stage about the performance participants in the first stage, conditional on the antecedents of Hypothesis 3.

5. Extensions and limitations

5.1. Other unrelated channels

Throughout the paper, we have focused on rewards as the main parameter that can possibly distort the relationship between expected accuracy and subjective complexity. The reason is that the reward is an experimental variable that can be perfectly observed and manipulated by the experimenter. However, there are other such parameters that are correlated with expected accuracy, but are not related in any way with subjective complexity.

These are individual parameters that enter the agent's optimization problem, such as risk preferences (captured by the utility function v), or skill (captured by the cost function c). It is not difficult to verify though that such parameters affect all tasks in essentially the same way, and thus will not lead to reversal in expected accuracy like reward will. That is formally, the probability of solving A will be larger than the probability of solving B (for a given reward) under some utility function and cost function, if and only if the same holds for some other utility function and cost function.

In the upcoming sections we further discuss distortions due to asymmetries and/or task-dependencies in these parameters.

5.2. Task-dependent utilities

Suppose that the agent's utility from answering correctly given some reward $x \in X$ is task-dependent, i.e., it is given by $v_S(x)$ rather than simply $v(x)$. We assume that the agent's risk preferences for money remain constant across tasks, i.e., we have

$$v_S(x) = \alpha_S + \beta_S v(x),$$

where $\alpha_S \geq 0$ and $\beta_S > 0$. We will henceforth write $\tilde{\alpha}_S := \alpha_S/\kappa_S$ and $\tilde{\beta}_S := \beta_S/\kappa_S$, as well as

$$\tilde{v}_S(x) := \frac{v_S(x)}{\kappa_S} = \tilde{\alpha}_S + \tilde{\beta}_S v(x).$$

Analogously to the benchmark case where utilities are task-independent, the agent's optimal attention strategy is determined by the threshold q_S^x, which is in turn the solution to equation $c'(q) = \min\{\tilde{v}_S(x), c'(0)\}$. Let $x_S := \sup\{x \geq 0 : q_S^x > 0\}$ denote the lowest reward that leads to full revelation of the state in task S. Notice that this threshold is such that

$$v(x_S) = \frac{c'(0) - \tilde{\alpha}_S}{\tilde{\beta}_S} = \frac{c'(0)\kappa_S - \alpha_S}{\beta_S}. \tag{13}$$

Obviously, if $c'(0) = \infty$ then $x_S = \infty$, like for instance when c is entropic.

For any two tasks $S, S \in \mathcal{S}$, let $\tilde{x}$ be the unique solution to $\tilde{v}_S(x) = \tilde{v}_{S'}(x)$ if $\tilde{\beta}_S \neq \tilde{\beta}_{S'}$, and set $\tilde{x} := \infty$ if $\tilde{\beta}_S = \tilde{\beta}_{S'}$. Then, define

$$\tilde{q} := \begin{cases} q_S^{\tilde{x}} = q_{S'}^{\tilde{x}} & \text{if } 0 < \tilde{x} < \min\{x_S, x_{S'}\}, \\ 1/2 & \text{otherwise,} \end{cases}$$

and subsequently take the corresponding degree of uncertainty

$$H(\tilde{q}) = -\tilde{q}\log\tilde{q} - (1-\tilde{q})\log(1-\tilde{q}). \tag{14}$$

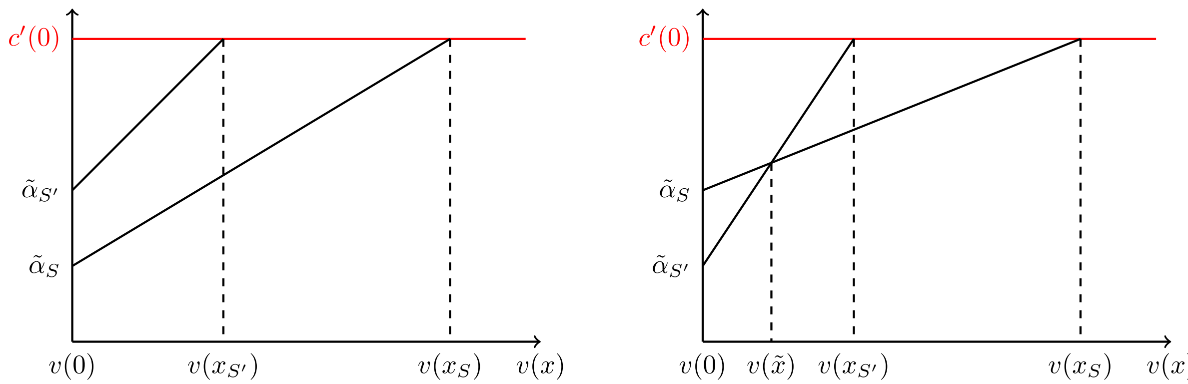
Obviously, whenever $\tilde{x}$ does not belong to the interval $\big(0, \min\{x_S, x_{S'}\}\big)$, we will have $H(\tilde{q}) = 1$. Before proceeding, note that whenever we write $\tilde{x}$ or $\tilde{q}$ we omit reference to pair S, S', even though both of them depend on the specific tasks. This is just for notation simplicity.

Before proceeding with the generalization of our results, let graphically illustrate the new elements that we just introduced (Figure 8).

Remark 3. It is not difficult to verify that the two tasks will not be $\succeq$-comparable whenever

$$\min\{H(\mu_S), H(\mu_{S'})\} > H(\tilde{q}). \tag{15}$$

This means that when there is high prior uncertainty about both of them, we will not be able to classify one as subjectively more complex than the other, even if they share the same prior and one is more difficult than the other. The intuition is that both for small and large rewards, there is a lot to learn in each of the tasks, as the prior uncertainty is high. At the same time, the utility from answering correctly is higher for one of the two tasks when the reward is small, and higher for the other task when the reward is large. And this is exactly what causes the reversal in the relation between $P(S, x)$ and $P(S', x)$ as x changes from small to large. For a formal proof of this statement, see Proposition B1 in Appendix B. $\triangleleft$



(a) $\tilde{x} < 0$: Thus, we have $\tilde{v}_{S'}(x) > \tilde{v}_S(x)$ for every $x \in (0, x_{S'})$, meaning that $q^x_{S'} \leq q^x_S$ for all $x \in X$. As a result, we obtain $H(\tilde{q}) = 1$. This case is qualitatively similar to the one with task-independent utilities.

(b) $0 < \tilde{x} < x_{S'}$: Thus, we have $\tilde{v}_{S'}(x) < \tilde{v}_S(x)$ and a fortiori $q^x_{S'} \geq q^x_S$ for every $x \in (0, \tilde{x})$, and $\tilde{v}_{S'}(x) > \tilde{v}_S(x)$ and a fortiori $q^x_{S'} \leq q^x_S$ for every $x \in (\tilde{x}, x_{S'})$. Hence, we obtain $H(\tilde{q}) \in (0, 1)$. This case differs from the one with task-independent utilities.

Figure 8: In each of the two cases, we depict $\tilde{v}_S(x)$ and $\tilde{v}_{S'}(x)$ as linear functions of $v(x)$. When facing task S the agent picks the perfectly informative signal if the reward is $x \geq x_S$, in which case the straight line that corresponds to $\tilde{v}_S$ goes above $c'(0)$. And likewise when facing S'.

We are now ready to define the two components that we will use in our characterization result. In particular, we define the vector-valued function $\tilde{\boldsymbol{\phi}} : \mathcal{S} \to \mathbb{R}^2$ by

$$\tilde{\boldsymbol{\phi}}_1(S) = v(x_S) \text{ and } \tilde{\boldsymbol{\phi}}_2(S) = \boldsymbol{\phi}_2(S). \tag{16}$$

Notice that $\tilde{\boldsymbol{\phi}}_1(S)$ can be seen as a generalization of $\boldsymbol{\phi}_1(S)$, in the sense that as long as $\alpha_S = \alpha_{S'}$ and $\beta_S = \beta_{S'}$, it will be the case that $\tilde{\boldsymbol{\phi}}_1(S) \geq \tilde{\boldsymbol{\phi}}_1(S')$ if and only if $\kappa_S \geq \kappa_{S'}$. Moreover, if $c'(0) = \infty$, consistently with (13), we will have $v(x_S) > v(x_{S'})$ if and only if either (i) $\tilde{\beta}_S < \tilde{\beta}_{S'}$, or (ii) $\tilde{\beta}_S = \tilde{\beta}_{S'}$ and $\tilde{\alpha}_S < \tilde{\alpha}_{S'}$. Then, we are ready to generalize our two main results to task-dependent utilities.

Proposition 2. *Fix a pair of tasks $S, S' \in \mathcal{S}$ such that $\min\{H(\mu_S), H(\mu_{S'})\} \leq H(\tilde{q})$. Then, the following hold:*

(i) $S \succeq S'$ if and only if $\tilde{\boldsymbol{\phi}}(S) \geq \tilde{\boldsymbol{\phi}}(S')$.

(ii) If $\tilde{\boldsymbol{\phi}}_1(S) > \tilde{\boldsymbol{\phi}}_1(S')$ and $\tilde{\boldsymbol{\phi}}_2(S) < \tilde{\boldsymbol{\phi}}_2(S')$, there exist $0 < \tilde{x}_1 < \tilde{x}_2 \leq \infty$ such that

(ii.1) $P(S, x) > P(S', x)$ for all $x \in (0, \tilde{x}_1)$,

(ii.2) $P(S, x) < P(S', x)$ for all $x \in (\tilde{x}_1, \tilde{x}_2)$,

(ii.3) $P(S, x) = P(S', x)$ for all $x \in (\tilde{x}_2, \infty)$.

5.3. State-dependent utilities

Suppose that the agent's utility function v_S is not only task-dependent but also state-dependent, i.e., the agent has (unobservable) stakes over the state realization in S (Tsakas, 2026, and references therein). This means that, for each reward $x \in X$, the utility of guessing correctly is equal to $v_S^k(x)$ when the correct answer is s_k. Still, we maintain the assumption that risk preferences are constant across states, i.e., we have

$$v_S^k := \alpha_S^k + \beta_S^k v_S,$$

where $\alpha_S^k, \beta_S^k > 0$. The fact that utility is state-dependent will induce distortions in expected accuracy that will lead to counterintuitive results. In particular, expected accuracy will not always be increasing in the reward or decreasing in pure difficulty, even if we keep prior uncertainty fixed.

Let us illustrate these results in the simple case where the cost is entropic (i.e., $c = H$), and the intrinsic incentives without any reward are state-independent (i.e., the utility parameters are $\alpha_S^0 = \alpha_S^1 = 0$ and $\beta_S^0 < \beta_S^1$). Then, the indirect utility function is given by

$$g_S^x(q) = v(x) \max\left\{\beta_S^0(1-q), \beta_S^1 q\right\}.$$

Thus, the belief that makes the agent indifferent between guessing s_0 and guessing s_1 will be $\bar{q}_S := \beta_S^0/(\beta_S^0 + \beta_S^1) < 1/2$. As a result, the posteriors that determine the attention region will now become

$$q_{s_0}^x = \frac{e^{\lambda_0 v(x)} - 1}{e^{(\lambda_0+\lambda_1)v(x)} - 1} \text{ and } q_{s_1}^x = \frac{e^{(\lambda_0+\lambda_1)v(x)} - e^{\lambda_1 v(x)}}{e^{(\lambda_0+\lambda_1)v(x)} - 1},$$

where $\lambda_0 := \beta_S^0/\kappa_S$ and $\lambda_1 := \beta_S^0/\kappa_S$. Standard algebra yields $q_{s_1}^x < 1/2$ for all $x \in X$ such that

$$v(x) < \frac{\kappa_S}{\beta_S^0} \log \frac{2\beta_S^1}{\beta_S^0 + \beta_S^1}. \tag{17}$$

For illustration purposes, throughout this section, assume that $v(0) = 0$. That is, for every task S for which utilities are state-dependent, there is some reward $\bar{x}_S > 0$ such that for all $x > \bar{x}_S$ expected accuracy $P(S, x)$ takes the form which is shown in Figure 9.

Now, take a task S with state-dependent utilities, together with a reward $x < \bar{x}_S$ and a prior belief $\mu_S \in (q_{s_0}^x, \bar{q}_S)$. Moreover consider some $x' < x$, which is sufficiently small so that the optimal signal is completely uninformative. Such x' will always exist due to the fact that $q_{s_0}^{x'}$ continuously approaches $\bar{q}_S$ from below as x' decreases, and eventually it will become larger than μ_S. This means that $P(S, x') > P(S, x)$, despite $x' < x$.

For similar reasons, subjective accuracy is sometimes increasing with respect to difficulty. In particular, take two tasks S and S' which are identical in every aspect other than difficulty, i.e., we have $\mu_S = \mu_{S'}$ and $v_S = v_{S'}$, but at the same time $\kappa_{S'} > \kappa_S$. Then, using a similar argument as the one above, we will obtain $P(S', x) > P(S, x)$, i.e., the agent has better chances at solving the more difficult task.

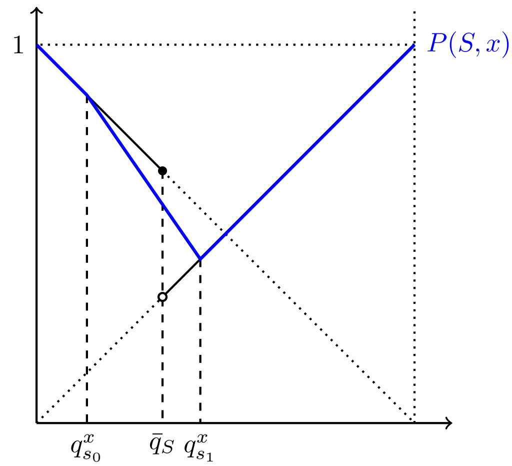


Figure 9: Given the task S and the reward x, the blue piecewise linear function is the expected accuracy (as a function of the agent's prior), assuming that the optimal signal π^x_S has been used.

The reason why state-dependent preferences yield such counterintuitive predictions is that agent's indirect utility $g^x(q)$ is not monotonic in the probability of being correct. This is because $P(S,x)$ is discontinuous at $\bar{q}_S$. Intuitively, because of the agent's intrinsic preferences for state s_1, she will not necessarily strive to maximize the probability of answering correctly. Hence, in the presence of stakes about the state space, expected accuracy is not guaranteed to be a good proxy for subjective complexity.

5.4. Elicitation of expected accuracy order

Of course, any practical use of the aforementioned results fully relies on our ability to elicit the uniform expected accuracy order, i.e., we want to know whether $P(S,x) \geq P(S',x)$ or $P(S,x) \leq P(S',x)$ for different $x \in X$. But unfortunately, direct elicitation of $P(S,x)$ is impossible, even if the utility function is both task-independent and state-independent. Let us explain why.

For each task $S \in \mathcal{S}_0$ define the state space $Z_S = \{0,1\}$, where 0 and 1 are respectively interpreted as the agent guessing wrongly and correctly in the task S. A belief elicitation mechanism yields a prize as a function of her reported probability $R(S,x)$ and the state realization in Z_S. Of course, the agent intrinsically cares about Z_S, viz., at state 1 she will receive a side payoff $x > 0$, whereas at state 0 she will receive 0. These side payoffs are on top of the prize that the belief elicitation mechanism yields. Hence, it is well-known from the literature on state-dependent SEU that $P(S,x)$ cannot be identified using traditional choice data, meaning that standard belief elicitation mechanisms would not elicit the agent's expected accuracy (Tsakas, 2026, and references therein).

At the same time, it is not $P(S,x)$ per se that we would like to elicit, but rather the relation between $P(S,x)$ and $P(S',x)$. This means that, even if we do not manage to truthfully elicit the agent's actual beliefs, it suffices to elicit how she ranks different tasks in terms of expected accuracy. As it turns out, this is always possible using a stochastic elicitation mechanism, e.g., a binarized scoring rule (Hossain and Okui, 2013).

Before presenting this result, let us recall how the binarized scoring rule works. For notation simplicity, whenever it is obvious from the context, we omit reference to the task S and the reward x, and we simply write $P := P(S, x)$ and $R := R(S, x)$ respectively. Then, based on how accurate R is in relation to the realized state in Z_S, she receives a number of lottery tickets for a fixed prize $y > 0$. In particular, the chances that she wins the prize are shown below:

	Wrong guess ($z = 0$)	Correct guess ($z = 1$)
R	$1 - \gamma R^2$	$1 - \gamma(1 - R)^2$

The parameter $\gamma \in (0, 1)$ determines the strength of the incentives provided by the scoring rule. It is without loss of generality to set $\gamma = 1$. As a result, the total probability of guessing wrongly and winning the prize is $(1 - P)(1 - R^2)$, and the corresponding outcome is $(0, y) \in Z_S \times \mathbb{R}$. Likewise, the total probability of guessing correctly and winning the prize is $P(1-(1-R)^2)$, and the corresponding outcome is $(1, x + y) \in Z_S \times \mathbb{R}$. And finally, the total probability of guessing correctly and losing the prize is $P(1 - R)^2$, and the corresponding outcome is $(1, x) \in Z_S \times \mathbb{R}$.

In this sense, by reporting R the agent chooses from a menu of acts. As usual, we assume that she will submit a report that maximizes her overall expected utility. In the standard setting, where the agent does not have any stakes in the state realization, the binarized scoring rule guarantees that P is the only optimal report, i.e., truth telling is strictly incentive compatible. Here, it is no longer the case, as there are side payoffs on top of the payment y that the scoring rule may yield. Nonetheless, the optimal reports suffice for eliciting the expected accuracy order.

Proposition 3 (Elicitation result)**.** *For every pair $S, S' \in \mathcal{S}$ and every $x \geq 0$:*

$$P(S, x) \geq P(S', x) \;\Leftrightarrow\; R(S, x) \geq R(S', x). \tag{18}$$

Despite the positive message given by the previous result, there are still two question marks regarding implementation. First, if we pay both for the actual task and for the elicitation mechanism, hedging opportunities will arise for the subject (Blanco *et al.*, 2010). Fortunately, this can be easily dealt with, by randomly paying only for one of the two. Second, there is an issue with the timing of elicitation. In particular, if $R(S, x)$ is reported after the agent has undertaken the task, like in Enke and Graeber (2023), we will no longer be able to vary the reward. Hence, the only option would be to elicit expected accuracy before the task is undertaken, using the strategy method across different rewards. The timing of elicitation is further discussed in Section 5.6.

An alternative approach that would deal with both aforementioned issues would be to rely on an idea similar to the one in Bayesian markets (Baillon, 2017), where we use beliefs about other individuals to proxy one's own beliefs about themselves. Accordingly, in our setting, we can elicit our agent's beliefs about the accuracy of another individual for different reward levels, under the assumption that the agent considers this other individual being similar to themselves. This is exactly the method we implement in our experiment in Section 4.

5.5. Asymmetric cost functions

Throughout the paper, we have assumed that the cost function c is symmetric. The main criticism that symmetry has received in the literature is based on the idea that distinguishing between two states might be more difficult than distinguishing between two other states (Hébert and Woodford, 2021). In other words, concerns regarding asymmetries should enter the picture primarily in cases where the state space has some underlying distance, and more similar states are harder to tell apart. However, given that throughout the paper we focus on binary tasks, it is well justified to maintain symmetric costs.

Nonetheless, we point out that symmetry is in fact crucial for our results. Let us illustrate this with a simple example that minimally deviates from the symmetric case. Suppose that the cost function is given by the weighted version of Shannon entropy

$$c(q) = -\lambda q \log q - (1-\lambda)(1-q)\log(1-q),$$

where $\lambda \in [0,1]$ describes the degree of asymmetry in the cost for acquiring information, with $\lambda = 1/2$ corresponding to the symmetric case. For the sake of presentation, we set $\lambda = 1$.

The crucial feature of this cost specification is that the two posteriors in the support of the optimal signal π_S^x are no longer symmetric around $1/2$. In particular, we have

$$q_{s_0}^x = \frac{v(x)}{\kappa_S\left(e^{2v(x)/\kappa_S} - 1\right)} \text{ and } q_{s_1}^x = \frac{v(x)e^{2v(x)/\kappa_S}}{\kappa_S\left(e^{2v(x)/\kappa_S} - 1\right)}.$$

For any fixed $x \in X$, if S is sufficiently easy, the ratio $v(x)/\kappa_S$ becomes large, in which case we obtain $q_{s_0}^x > 1 - q_{s_1}^x$. This means that subjective accuracy as a function of the prior belief will be tilted like in Figure 10.

Then, take another task $S' \in \mathcal{S}$ which is easier than S, i.e., $\kappa_{S'} < \kappa_S$. This will directly imply $q_{s_0'}^x < q_{s_0}^x$ and $q_{s_1'}^x > q_{s_1}^x$. Nonetheless, if the difference in difficulty is small (i.e., if $\kappa_{S'}$ is not much smaller than κ_S), it will be the case that $q_{s_0'}^x > 1 - q_{s_1}^x$. Hence, if take $\mu_{S'}$ close to $q_{s_0'}^x$ and μ_S in the interval $(1 - q_{s_0'}^x, q_{s_1}^x)$, it will be the case that $P(S,x) > P(S',x)$, despite the fact that S is both more difficult and less known a priori than S'.

Such counterexamples can be always constructed when the cost function is asymmetric, implying that symmetry is necessary for subjective accuracy to be a good proxy of complexity. However, as we have already mentioned above, this is not a major concern, at least when it comes to binary tasks, like the ones we consider here.

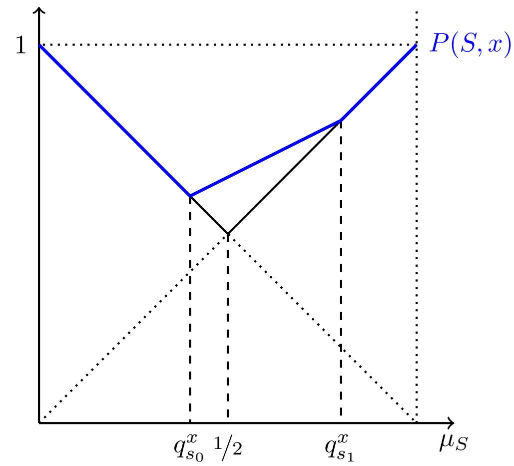


Figure 10: Given the task S and the reward x, the blue piecewise linear function is the expected accuracy (as a function of the agent's prior), assuming that the optimal signal π^x_S has been used.

5.6. Ex ante vs ex post subjective accuracy

Throughout the paper we have focused on ex ante subjective accuracy, i.e., $P(S,x)$ is the probability that the agent attaches to eventually answering correctly in task S before the signal π^x_S is actually realized, and a fortiori before a guess is actually submitted. This is in contrast to previous work on cognitive uncertainty where subjects were asked ex post how confident they were on their previously submitted answer being correct (Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Agranov *et al.*, 2025).

However, it is not difficult to see that in our setting $P(S,x)$ is also equal to the ex post subjective accuracy. In particular, if $\mu_S \in (q^x_S, 1-q^x_S)$, one of the two posterior beliefs in $\text{supp}(\pi^x_S) = \{q^x_S, 1-q^x_S\}$ will be eventually realized. But then, under both posteriors the ex post subjective accuracy will be equal to $1-q^x_S$, which is equal to $P(S,x)$ itself. On the other hand, if $\mu_S \notin (q^x_S, 1-q^x_S)$, the optimal signal will trivially attach probability 1 to μ_S itself, meaning that ex ante and ex post subjective accuracy will again be equal to $P(S,x)$.

The previous observation suggests that whether we elicit expected accuracy before or after a guess has been made will not make any difference in our measurement of $P(S,x)$. Nevertheless, as we have already mentioned in Section 5.4, ex post expected accuracy can only be elicited for a single reward, as opposed to ex ante expected accuracy which can be elicited for all rewards using the strategy method.

5.7. Exogenous information

What primarily distinguishes our measure from other notions of subjective complexity is the role of prior uncertainty. As we have already discussed, differences in prior uncertainty reflect differences in prior information about the tasks. Thus, since prior (exogenous) information has rendered the tasks incomparable, new (exogenous) information can also restore comparability. Interestingly, this will be the case without needing to assume much about the underlying information stream.

The term "exogenous" refers to information that is not generated by the signal that the agent implements as part of her attention strategy. Formally, this takes the form of a sequence of iid draws from a data-generating process which is correlated with a task S. The agent may have misspecified beliefs on the data generating process, like in the literature on misspecified learning. Yet, it follows directly from standard results that the agent's belief about S will converge to a Dirac measure almost surely (e.g., Berk, 1966; Shalizi, 2009).

As a result, if $\kappa_{S'} > \kappa_S$, information cannot change the fact that S' is more difficult than S, and therefore S will never be deemed more complex than S' regardless of whether the information is about S or S'. On the other hand, information about S will always lead to beliefs such that S is classified as more complex than S'.

5.8. Monotonicity in rewards

In their recent paper Alaoui and Penta (2022) discuss the fact that rewards are sometimes detrimental to performance, with performance in our model being taken as a synonym for expected accuracy. They cite different mechanisms that can potentially lead to such a paradoxical result.

For example, we may have cost-based explanations, which in our setting translate to the cost function increasing in the reward, and therefore higher x leading to a less informative π^x_S. This is for instance the case when higher incentives cause more pressure. Alternatively, we may also have value-based explanations, which in our setting translate to $v(x)$ being not necessarily increasing in x, and again higher x leading to a less informative π^x_S. Examples of such phenomena include distractions that diminishes motivation.

In all of these cases, our single-crossing result (Theorem 2) will typically no longer hold. On the other hand, whether our characterization result (Theorem 1) will still hold depends on the exact specification. For example, even if v is not increasing, as long as it is unbounded, our result will still hold.

Finally, let us point out that earlier (Section 5.3) we have also identified another possible mechanism that may render additional rewards detrimental to expected accuracy, viz., larger reward may lead to smaller expected accuracy ceteris paribus.

6. Conclusion

In this paper, we proposed uniform expected accuracy as a measure of subjective complexity. This measure maintains the appealing features of expected accuracy, which has recently become a standard measure of complexity (Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Agranov *et al.*, 2025). In particular, this measure is simple, intuitive and easily portable across tasks.

Moreover, it captures the two main channels that connect expected accuracy with subjective complexity, viz., pure difficulty and prior uncertainty. But at the same time, it also controls for the main exogenous parameter that affects expected accuracy but is not related to subjective complexity, viz., the reward.

Then, we proceed to characterize the incomplete order over tasks, which is induced by uniform expected accuracy. Our characterization result implies that task A is classified as subjectively more complex than task B if and only if A is both more difficult and much less known than B. In this sense, we provide theoretical foundations to the intuitive idea that complexity has both an objective and a subjective part (Oprea, 2024a). This conclusion is also aligned with the literature in psychology (Campbell, 1988) and information science (Vakkari, 2005) which —unlike economics— have identified prior uncertainty as a key dimension of complexity.

Finally, using a lab experiment, where we can exogenously control both difficulty and uncertainty, we corroborate our theoretical predictions. This suggests that our work supports the recently surging experimental literature that uses expected accuracy as a proxy for subjective complexity (Enke and Graeber, 2023; Enke *et al.*, 2024a,b; Agranov *et al.*, 2025), under the condition that expected accuracy is elicited for multiple different rewards using the strategy method.

A. Proofs

Proof of Proposition 1. By a standard concavification argument, there exists some $q_S^x \in [0, 1/2]$ such that the optimal signal π_S^x distributes all probability between q_S^x and $1-q_S^x$ whenever $q_S^x < \mu_S < 1 - q_S^x$, and it is completely uninformative otherwise. Define

$$\kappa := \sup\left\{\kappa_S \;\middle|\; S \in \mathcal{S}_0 \text{ such that } q_S^0 = 0\right\}. \tag{A.1}$$

$(\Rightarrow)$: Take an arbitrary $S \in \mathcal{S}_0$. First, if $H(\mu_S) = 0$, then $P(S, x) = 1$ for all $x \geq 0$. Second, if $\kappa_S \leq \kappa$ then $q_S^0 = 0$,which together with $q_S^x \leq q_S^0$ implies $q_S^x = 0$, and therefore $P(S, x) = 1$ for all $x \geq 0$. Putting the two together completes this part of the proof.

$(\Leftarrow)$: For an arbitrary $S \in \mathcal{S}_0$, suppose that $\kappa_S > \kappa$ and $H(\mu_S) > 0$. Then, by (A.1), we have $q_S^0 > 0$, and therefore $P(S, 0) < 1$, which completes the second part of the proof. □

Proof of Theorem 1. $(\Leftarrow)$: By $\min\{H(q_S^0), H(\mu_S)\} \geq \min\{H(q_{S'}^0), H(\mu_{S'})\}$, we obtain $H(\mu_S) \geq \min\{H(q_{S'}^0), H(\mu_{S'})\}$. Then, consider two cases:

1) $H(\mu_S) \geq H(\mu_{S'})$: This means $\max\{\mu_S, 1-\mu_S\} \leq \max\{\mu_{S'}, 1-\mu_{S'}\}$. Moreover, by q_S^x being increasing in κ_S, we have $1 - q_S^x \leq 1 - q_{S'}^x$. Hence, by Equation (8), we obtain $P(S, x) \leq P(S', x)$ for all $x \geq 0$.
2) $H(\mu_S) < H(\mu_{S'})$: This means that $H(\mu_S) \geq H(q_{S'}^0)$, and a fortiori $\max\{\mu_S, 1-\mu_S\} \leq 1 - q_{S'}^0$. Moreover, by $q_{S'}^x$ being decreasing in x, we have $1 - q_{S'}^x \geq 1 - q_{S'}^0$. Hence, we

obtain $\max\{\mu_S, 1-\mu_S\} \leq 1 - q^x_{S'}$. Then, together with $1 - q^x_S \leq 1 - q^x_{S'}$, we obtain

$$\max\{\mu_S, 1-\mu_S, 1-q^x_S\} \leq 1 - q^x_{S'} \leq \max\{\mu_{S'}, 1-\mu_{S'}, 1-q^x_{S'}\},$$

which, by Equation (8), implies $P(S,x) \leq P(S',x)$ for all $x \geq 0$.

Putting the two cases together directly yields $S \succeq S'$.

$(\Rightarrow)$: Sufficiency follows directly from Theorem 2, which we prove below. □

Proof of Theorem 2. By $\kappa_S = \boldsymbol{\phi}_1(S) > \boldsymbol{\phi}_1(S') = \kappa_{S'}$, we obtain for all $x \geq 0$

$$q^x_{S'} \leq q^x_S \tag{A.2}$$

with equality holding if and only if $q^x_S = 0$. Applying the latter for $x = 0$, together with the fact that S and S' are non-trivial, yields

$$H(q^0_{S'}) < H(q^0_S). \tag{A.3}$$

By combining (A.3) with $\min\{H(\mu_S), H(q^0_S)\} = \boldsymbol{\phi}_2(S) < \boldsymbol{\phi}_2(S') = \min\{H(\mu_{S'}), H(q^0_{S'})\}$, we obtain the following two inequalities:

$$H(\mu_S) < H(\mu_{S'}), \tag{A.4}$$
$$H(\mu_S) < H(q^0_{S'}). \tag{A.5}$$

By $q^x_{S'}$ being strictly decreasing in x, there exists a unique $x_1 > 0$ such that $\mu_S \in (0, q^x_{S'}) \cup (1 - q^x_{S'}, 1)$ for all $x < x_1$. Thus, by combining (A.4) and (A.5) with (8), we obtain $P(S,x) > P(S',x)$. On the other hand, for all $x > x_1$, we have $\mu_S \in (q^x_{S'}, 1 - q^x_{S'})$. Hence, by (A.4), (A.5) and (8), we obtain $P(S,x) \leq P(S',x)$, with equality holding if and only if $x \geq x_2 := \inf\{x > 0 : P(S,x) = 1\}$. The latter completes the proof. □

Proof of Proposition 2. *(i)* $(\Leftarrow)$: By $\tilde{\boldsymbol{\phi}}_1(S) \geq \tilde{\boldsymbol{\phi}}_1(S')$ we obtain $x_S \geq x_{S'}$. This is because v is strictly increasing in X. Then, we distinguish two possible cases:

1) $\tilde{\alpha}_S \leq \tilde{\alpha}_{S'}$: This is the case illustrated in Figure 8.a, where $q^x_S \geq q^x_{S'}$ for all $x \in X$. Thus, the remainder of the proof is identical to the one of Theorem 1.

2) $\tilde{\alpha}_S > \tilde{\alpha}_{S'}$: This is the case illustrated in Figure 8.b. Let us first focus on the interval $x \in [0, \tilde{x})$, where $\tilde{v}_S(x) > \tilde{v}_{S'}(x)$ and a fortiori $q^x_S < q^x_{S'}$. By setting the reward to $x = 0$, we obtain $\tilde{q} < q^0_S < q^0_{S'}$, and therefore $H(\tilde{q}) < H(q^0_S) < H(q^0_{S'})$. Moreover, by $\tilde{\boldsymbol{\phi}}_2(S) \geq \tilde{\boldsymbol{\phi}}_2(S')$, we have $H(\mu_S) \geq \min\{H(\mu_S), H(q^0_S)\} \geq \min\{H(\mu_{S'}), H(q^0_{S'})\}$.

   Suppose $\min\{H(\mu_{S'}), H(q^0_{S'})\} = H(q^0_{S'})$. This would in turn imply $H(\tilde{q}) < H(q^0_{S'}) \leq \min\{H(\mu_S), H(\mu_{S'})\}$, which contradicts our hypothesis that $H(\tilde{q}) \geq \min\{H(\mu_S), H(\mu_{S'})\}$. So, it must necessarily be the case that $\min\{H(\mu_{S'}), H(q^0_{S'})\} = H(\mu_{S'})$, meaning that

$H(\mu_S) \geq H(\mu_{S'})$. By $H(\tilde{q}) \geq \min\{H(\mu_S), H(\mu_{S'})\}$, we also have $H(\tilde{q}) \geq H(\mu_{S'})$. Hence, by (8) we conclude $P(S', x) > P(S, x)$ for all $x < \tilde{x}$.

Then, let us turn focus on the interval $x \in [\tilde{x}, \infty]$, where $\tilde{v}_S(x) \leq \tilde{v}_{S'}(x)$ and a fortiori $q_S^x \geq q_{S'}^x$. As we have already established that $\eta_{S'} \leq \eta_S$, it follows again by (8) that $P(S', x) \geq P(S, x)$ for all $x \geq \tilde{x}$.

Putting the cases two together completes the proof of sufficiency.

(i) $(\Rightarrow)$: By $\tilde{\boldsymbol{\phi}}_1(S) > \tilde{\boldsymbol{\phi}}_1(S')$, we have $x_S > x_{S'}$. Once again, consider the same two cases we had in the proof of sufficiency:

1) $\tilde{\alpha}_S \leq \tilde{\alpha}_{S'}$: It is the case that $q_S^x \geq q_{S'}^x$ for all $x \in X$, and the proof follows the same steps as the one of Theorem 2.

2) $\tilde{\alpha}_S > \tilde{\alpha}_{S'}$: Suppose that $\eta_{S'} > \tilde{\eta}$. Then, there is some $x < \tilde{x}$ such that $H(\tilde{q}) < H(q_S^x) < H(q_{S'}^x) < H(\mu_{S'})$. Hence, by (8), we obtain $P(S, x) = P(S', x)$, which contradicts $S \succeq S'$. So, it must be the case that $H(\mu_{S'}) \leq H(\tilde{q})$. Thus, we have

$$H(\mu_{S'}) = \min\{H(\mu_{S'}), H(q_{S'}^0)\} > \min\{H(\mu_S), H(q_S^0)\} = H(\mu_S),$$

where the first and last equality follow from $H(\tilde{q}) < H(q_S^0) < H(q_{S'}^0)$, and the inequality in the middle follow from $\tilde{\boldsymbol{\phi}}_2(S) < \tilde{\boldsymbol{\phi}}_2(S')$. Hence, for every $x < \tilde{x}$ it is the case that $P(S, x) > P(S', x)$. Moreover, using the same argument as the one in Theorem 2, there exists some $\tilde{x}_2 > \tilde{x}_1 > \tilde{x}$ such that $P(S, x) > P(S', x)$ for $x \in (\tilde{x}, \tilde{x}_1)$, whereas $P(S, x) < P(S', x)$ for $x \in (\tilde{x}_1, \tilde{x}_2)$.

Putting the cases two together completes the proof of necessity.

(ii) : It has been already proven in the necessity part above. □

Proof of Proposition 3. Let $u : Z_S \times \mathbb{R} \to \mathbb{R}$ be the utility function, written in each explicit form. This means that our previously-defined net utility is simply $v(x) = u(1, x) - u(0, 0)$. So, for task $S \in \mathcal{S}$ and reward $x \geq 0$, the overall expected utility from reporting R is equal to

$$\mathbb{E}_P(u(R)) = (1 - P)(1 - R^2)u(0, y) + P(1 - (1 - R)^2)u(1, x + y) + P(1 - R)^2 u(1, x).$$

Then, the first order condition in the choice variable R yields

$$-(1 - P)Ru(0, y) + P(1 - R)u(1, x + y) - P(1 - R)u(1, x) = 0,$$

which is in turn equivalent to

$$\frac{1 - R}{R} = \frac{1 - P}{P} \cdot \frac{u(0, y)}{u(1, x + y) - u(1, x)}.$$

Since the second fraction on the righthand side is task independent and strictly positive, it follows that R is strictly increasing in P. □

B. Additional results

Proposition B1. *Consider two tasks $S, S' \in \mathcal{S}$ with task-dependent utilities for solving them correctly. Moreover, let $\min\{H(\mu_S), H(\mu_{S'})\} > H(\tilde{q})$. Then, S and S' are not $\succeq$-comparable, i.e., we have neither $S \succeq S'$ nor $S' \succeq S$.*

Proof. First of all, this condition is trivially violated if $H(\tilde{q}) = 1$, meaning that it is only relevant in cases where $0 < \tilde{x} < \min\{x_S, x_{S'}\}$, like for instance in Figure 8.b. Then, by (15), there exists some $x < \tilde{x}$ such that $H(\tilde{q}) < H(q^x_{S'}) < H(q^x_S) < \min\{H(\mu_S), H(\mu_{S'})\}$, and therefore

$$P(S, x) = 1 - q^x_S < 1 - q^x_{S'} = P(S', x).$$

On the other hand, for any $x > \tilde{x}$, we obtain $H(q^x_S) < H(q^x_{S'}) < H(\tilde{q}) < \min\{H(\mu_S), H(\mu_{S'})\}$, and therefore

$$P(S, x) = 1 - q^x_S > 1 - q^x_{S'} = P(S', x).$$

This implies that neither $S \succeq S'$ nor $S' \succeq S$. Therefore, we will henceforth consider pairs of tasks such that $\min\{H(\mu_S), H(\mu_{S'})\} \leq H(\tilde{q})$. □

C. Experimental results

For the first stage of the experiment we invited 20 subjects who were randomly placed in the two treatments (6 participants in the low-stakes and 14 participants in the high-stakes treatment).[13] The descriptive statistics of the first stage looked as follows. Obviously, with the number of observations that we have, it would be meaningless to do any kind of statistical analysis. Thus, as explained in the main body of the paper, we use the first stage only as an auxiliary treatment, which is run with the sole purpose of not deceiving participants in the second stage.

	Low Stakes	**High Stakes**
EL	66.7	92.9
EH	50.0	100.0
DL	16.7	0.0
DH	50.0	42.9

Table 1: Percentage of correct answers for each task and each treatment.

For the second stage, we preregistered 100 participants. In the end, we only managed to recruit 57 subjects (30 females/27 males) from our pool. The experiment was conducted in May 2025 in the Behavioral and Experimental Economic Laboratory (BEELab) in Maastricht University.

[13]Due to random assignment the two treatments were not balanced. This is not a concern for our experiment as the first stage was merely used to calculate the payments in the second stage without deceiving the participants of the second stage.

Their average guesses for the percentage of first-stage participants that answered correctly are depicted in Table 2. This data is depicted graphically in Figure 6.

	Mean	**St. Dev.**
$P(EL, €10)$	82.6	15.7
$P(EL, €0.5)$	72.3	19.8
$P(EH, €10)$	72.6	18.9
$P(EH, €0.5)$	61.9	18.1
$P(DL, €10)$	68.9	19.5
$P(DL, €0.5)$	56.1	21.6
$P(DH, €10)$	59.7	16.5
$P(DH, €0.5)$	48.0	13.7

Table 2: The average guess for the expected accuracy of first-stage participants for each task and each reward.

In order to test Hypothesis 1, we ran a Wilcoxon signed rank test for each of the four tasks. The results are shown in Table 3.

Hypothesis	z	$\boldsymbol{Prob > \lvert z\rvert}$	**Exact Prob**
$P(EL, €10) = P(EL, €0.5)$	4.662	0.0000	0.0000
$P(EH, €10) = P(EH, €0.5)$	4.968	0.0000	0.0000
$P(DL, €10) = P(DL, €0.5)$	5.008	0.0000	0.0000
$P(DH, €10) = P(DH, €0.5)$	5.569	0.0000	0.0000

Table 3: The results of the Wilcoxon signed rank test for Hypothesis 1.

In order to test Hypothesis 2, we ran a Wilcoxon signed rank test for each inequality and for each reward. The results are shown in two separate tables, viz., in Table 4 for $H.2a$ and Table 5 for $H.2b$.

Hypothesis	z	$\boldsymbol{Prob > \lvert z\rvert}$	**Exact Prob**
$P(EL, €10) = P(EH, €10)$	5.438	0.0000	0.0000
$P(EL, €0.5) = P(EH, €0.5)$	4.650	0.0000	0.0000
$P(EH, €10) = P(DH, €10)$	5.370	0.0000	0.0000
$P(EH, €0.5) = P(DH, €0.5)$	5.244	0.0000	0.0000

Table 4: The results of the Wilcoxon signed rank test for Hypothesis $H.2a$.

In order to test Hypothesis 3, we ran two Wilcoxon signed rank tests. For Hypothesis $H.3a$, the number of participants who guessed $P(EH, 0.5) \geq P(DL, 0.5)$ was 32. For Hypothesis

Hypothesis	z	$Prob > \|z\|$	**Exact Prob**
$P(EL, €10) = P(DL, €10)$	4.534	0.0000	0.0000
$P(EL, €0.5) = P(DL, €0.5)$	4.621	0.0000	0.0000
$P(DL, €10) = P(DH, €10)$	3.714	0.0002	0.0001
$P(DL, €0.5) = P(DH, €0.5)$	3.171	0.0015	0.0012

Table 5: The results of the Wilcoxon signed rank test for Hypothesis $H.2b$.

$H.3b$, the number of participants who guessed $P(EH, 0.5) \geq P(DL, 0.5)$ was 28. The results are shown below in two separate tables, viz., in Table 6 for $H.3a$ and Table 7 for $H.3b$.

Hypothesis	z	$Prob > \|z\|$	**Exact Prob**
$P(EH, €10) = P(DL, €10)$	3.696	0.0002	0.0001

Table 6: The results of the Wilcoxon signed rank test for Hypothesis $H.3a$. That is, the hypothesis is tested conditional on $P(EH, €0.5) \geq P(DL, €0.5)$.

Hypothesis	z	$Prob > \|z\|$	**Exact Prob**
$P(EH, €0.5) = P(DL, €0.5)$	-2.186	0.0288	0.0280

Table 7: The results of the Wilcoxon signed rank test for Hypothesis $H.3b$. That is, the hypothesis is tested conditional on $P(EH, €10) \leq P(DL, €10)$.

D. Experimental instructions

The complete set of instructions for the first stage of the experiment can be found by clicking on this link and for the second stage on this link.

For the first stage, these are the instructions for the high stakes treatment. The only difference in the low stakes treatment is that the bonus of €10 is replaced by €0.50.

References

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

Agranov, M., Schotter, A. and Trevino, I. (2025). Complex for whom? an experimental approach to subjective complexity. *Working Paper*.

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

Avoyan, A., Ribeiro, M., Schotter, A., Schotter, E., Vaziri, M. and Zou, M. (2023). Planned vs. actual attention. *Management Science*, **70**, 2705–3380.

— and Schotter, A. (2020). Attention in games: An experimental study. *European Economic Review*, **124**, 103410.

Aydogan, I., Berger, L. and Théroude, V. (2023). More ambiguous or more complex? an investigation of individual preferences under uncertainty. *Working Paper*.

Baillon, A. (2017). Bayesian markets to elicit private information. *Proceedings of the National Academy of Sciences*, **114**, 7958–7962.

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

Berk, R. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Annals of Mathematical Statistics*, **37**, 51–58.

Blanco, M., Engelmann, D., Koch, A. and Normann, H.-T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, **13**, 412–438.

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

Campbell, D. (1988). Task complexity: a review and analysis. *The Academy of Management Review*, **13**, 40–52.

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

—, 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.

Danz, D., Vesterlund, L. and Wilson, A. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review*, **112**, 2851–2883.

de Clippel, G., Moscariello, P., Ortoleva, P. and Rozen, K. (2025). Caution in the face of complexity. *Working Paper*.

De Oliveira, H. and Mensch, J. (2025). Posterior-separable costs and menu preferences. *Working Paper*.

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*.

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

—, — and Oprea, R. (2024a). Complexity and time. *Journal of the European Economic Association*.

—, —, — and Yang, J. (2024b). Behavioral attenuation. *Working Paper*.

— and Shubatt, C. (2024). Quantifying lottery choice complexity. *Working Paper*.

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

Fudenberg, D. and Puri, I. (2022). Simplicity and probability weighting in choice under risk. *American Economic Review: Papers and Proceedings*, **112**, 421–425.

—, Strack, P. and Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, **108**, 3651–3684.

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

Gill, D. and Prowse, V. (2023). Strategic complexity and the value of thinking. *Economic Journal*, **133**, 761–786.

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

—, Nunnari, S. and Zarate-Pina, J. (2024). Revealing complexity. *Extended Abstract*.

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

Hossain, T. and Okui, R. (2013). The binarized scoring rule. *Review of Economic Studies*, **80**, 984–1001.

Hua Hu, E. (2023). A procedural model of complexity under risk. *Working Paper*.

Huck, S. and Weizsäcker, G. (1999). Risk, complexity, and deviations from expected-value maximization: Results of a lottery choice experiment. *Journal of Economic Psychology*, **20**, 699–715.

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.

Mononen, L. (2025). On preference for simplicity and probability weighting. *Working Paper*.

Mussolf, R. and Zimmermann, F. (2025). Model uncertainty. *Working Paper*.

Nagel, L. and Saitto, R. (2025). A measure of complexity for strategy-proof mechanisms. *Working Paper*.

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.

— (2024a). Complexity and its measurement. *Handbook of Experimental Methods in the Social Sciences*.

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

Puri, I. (2025). Simplicity and risk. *Journal of Finance*, **80**, 1029–1080.

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

Schlag, K. and van der Weele, J. (2013). Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality. *Theoretical Economics Letters*, **3**, 38–42.

Shalizi, C. R. (2009). Dynamics of bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics*, **3**, 1039–1074.

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

Shubbat, C. and Yang, J. (2024). Tradeoffs and comparison complexity. *Working Paper*.

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

Tsakas, E. (2020). Robust scoring rules. *Theoretical Economics*, **15**, 955–987.

— (2026). Belief identification by proxy. *Review of Economic Studies*, **93**, 697–724.

Vakkari, P. (2005). Task-based information searching. *Annual Review of Information Science and Technology*, **37**, 413–464.

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**.