Unpacking Overconfident Behavior

Abstract

We split overconfident behavior into two discrepancies from Bayesian behavior: a bias in beliefs, usually known as overconfidence, and attitudes towards ambiguity. We propose a method to separate these effects and apply it in two incentivized experiments. The first study reveals the importance of ambiguity attitudes in overconfident behavior. While beliefs were underconfident, subjects were ambiguity seeking, counterbalancing the effect of beliefs and leading to neither over- nor underconfident behavior. The second study applies the method in a setting where the amount of overconfident behavior is expected to vary: i.e., hard and easy questions. Previous studies suggested that for absolute performance people are overconfident for hard questions and underconfident for easy questions, but for relative performance the pattern is reversed. We show that this *hard-easy effect* for beliefs is robust to discrepancies from Bayesian choice behavior and incentivized decisions. On the other hand, easier tasks lead to more optimistic ambiguity attitudes for both absolute and relative performance.

Keywords: Overconfidence, subjective expected utility, belief measurement, ambiguity attitude, hard-easy effect.
JEL classification: D81, D83, D91.

1. Introduction

Behavioral research on overconfidence has traditionally focused on discrepancies from Bayesian behavior related to biases in beliefs when betting on performance, whether absolute or relative. It shows in particular that people tend to have too much confidence in their abilities. This overconfidence is one of the crown jewels of behavioral research and is, according to prominent scholars, the main cognitive bias (De Bondt and Thaler 1995, Kahneman 2011). It can be harmful and has been linked to entrepreneurial failure (Camerer and Lovallo 1999, Hooshangi and Loewenstein 2018), financial crises (Ho et al. 2016), undesirable business takeovers (Malmendier and Tate 2008), biased product selection (Feiler and Tong 2021), and political extremism (Ortoleva and Snowberg 2015).[1]

One challenge in the literature on overconfidence is how to measure beliefs. Scholars have increasingly used methods based on revealed preferences and observed behavior to measure overconfidence. For instance, CEOs' beliefs about the future performance of their firm have been inferred from the time at which they choose to exercise their stock options (Malmendier and Tate 2005). More overconfident CEOs exercise their options later because they have rosier beliefs about their performance, which in turn could affect corporate decisions (Malmendier and Tate 2008, Galasso and Simcoe 2011). Similarly, scholars have used methods such as proper scoring rules (Moore and Healy 2008) or matching probabilities (Mobius et al. 2014, Bruhin et al. 2018) to infer subjects' beliefs in the lab. Both in the field and the lab, these measures of beliefs assume a Bayesian choice setup (i.e., subjective expected utility). However, behavioral decision research has shown that, when inferred from choices, beliefs may be impacted by attitudes towards ambiguity, leading to attitudinal discrepancies from Bayesian behavior, in addition to the overconfidence bias in beliefs.[2] Descriptive decision theories have been proposed to generalize the classical Bayesian choice framework, where willingness to bet coefficients are substituted for subjective probabilities to allow for ambiguity attitudes. Consequently, the joint impact of overconfident beliefs and ambiguity seeking could result in overconfident behavior revealed through an inflated willingness to bet on winning events.

Our paper aims to unpack these two distinct causes of overconfident behavior. We think this is important, not only for organizations and governments that seek to reduce the harmful effects of overconfident behavior (and are more likely to succeed if they target the right cause) but also for prescriptive decision analysis, which aims to help people make better decisions. Overconfident biased probabilities are generally undesirable and need correction, but the prescriptive status of ambiguity aversion is less clear.[3] If ambiguity aversion and seeking are not considered desirable for prescriptive

[1] Moore and Healy (2008) distinguish three types of overconfidence: overestimation (thinking that you are better than you are), overplacement (thinking that you are better than others), and overprecision (thinking that outcomes are more certain than they are). In this paper, we study the first two types.

[2] Scoring rules, for example, are biased due to a non-neutral attitude towards risk (as revealed through utility under expected utility) and behavioral discrepancies from the expected utility, such as probability weighting (Winkler and Murphy 1970, Offerman et al. 2009).

[3] We note that a stream of the literature suggests that when performance is not fully correlated with true skills or confidence,

purposes, as is standard in decision analysis, they should be factored out.[4] Moreover, the appropriate correction depends on what causes overconfident behavior. Biases in beliefs may be reduced through statistical information (McGraw et al. 2004), advice, and mentorship (Bryan et al. 2017), but these measures may be less effective in reducing optimistic attitudes (for more on debiasing techniques, see Larrick 2004 and Soll et al. 2016).[5]

We used a choice-based method to measure subjective probabilities and ambiguity attitudes jointly and, thus, to assess their different effects on overconfident behavior. In terms of the ability to capture non-neutral ambiguity attitudes, our method accords with the main theories of decision under uncertainty, including α-maxmin and prospect theory, which do well empirically (Baillon and Bleichrodt 2015). We applied our method in two experiments in which subjects bet on their performance on an ability test. We used real incentives and, in line with recent recommendations, elicited subjective probability distributions rather than single probabilities (Soll et al. 2021). Subjects bet on both their absolute and relative performance on an ability test, allowing us to study both overestimation and overplacement as related to overconfident behavior.

The first study showed the importance of ambiguity attitudes for overconfident behavior. Analyses using uncorrected beliefs in a Bayesian choice framework suggested neither over- nor underconfidence. Instead, splitting willingness to bet, the indicator of overconfident behavior in our non-Bayesian choice setup, into subjective probabilities and ambiguity attitudes showed an interesting nuance. While subjective probabilities were underconfident as regards the uncertainty surrounding performance on the test, this was partly offset by subjects' ambiguity seeking. This ambiguity seeking (when betting on one's performance) was stronger than what has usually been found for exogenous sources of uncertainty like Ellsberg (1961) urns or the stock market's performance.

In a second study, we applied our method in a setting where the amount of overconfident behavior was expected to vary. Specifically, we investigated domains where over- and underconfidence are usually observed in Bayesian-like setups: easy and hard questions. The literature suggests a *hard-easy effect*: underestimation and overplacement for easy tasks and overestimation and underplacement for hard tasks (Moore and Healy 2008). We showed the robustness of the hard-easy effect in a non-Bayesian and incentivized choice setup. This challenges the claim that this effect disappears when offering incentives (Murad et al. 2016, Grieco and Hogarth 2009).

The second study also included an Ellsberg urn task, which allowed us to confirm that our endogenous source of uncertainty (betting on one's performance) leads to more ambiguity seeking than an exogenous source of uncertainty (betting on Ellsberg urns). As in the first study, we observed that

overconfidence in beliefs may not be a bias (e.g., Soll 1996, Glaser et al. 2013).

[4]For arguments that ambiguity aversion is not prescriptive see Raiffa (1961) and Wakker (2010). For arguments that it can be see Berger et al. (2017), Gilboa et al. (2011), and Gilboa and Marinacci (2016).

[5]The concept of attitudinal optimism we use in the current paper is specific to our non-Bayesian choice framework (e.g., Dimmock et al. 2016).

ambiguity attitudes played an essential role in overconfident behavior. Notably, we found that the task's difficulty affected not only beliefs but also attitudes. For both absolute and relative performance, subjects displayed more optimistic attitudes for the easy than for the hard task. These strong optimistic attitudes sometimes counterbalanced the pessimistic beliefs (such as for absolute performance on easy tasks) and sometimes exacerbated the optimistic beliefs (such as for relative performance on easy tasks).

2. Theoretical Framework

2.1. Notation and Definitions

In our experiments, subjects' performance on an ability test was ambiguous, i.e., uncertain with unknown probabilities. As subjects bet on performance on the test, the nature of uncertainty in our setup was epistemic as opposed to aleatory (Fox and Ülkümen 2011). We used a *state space*, which consisted of subjects' possible score (from 0% to 100% of correct answers) or rank (from 0% to 100%), to describe this ambiguity.

*Event*s are subsets of the state space. The notation $[s_\star, s^\star]$ means that a score [rank] was between $s_\star$% and $s^\star$%. We use $x_E y$ to denote the *ambiguous prospect* that pays x if event E occurs and y otherwise. If probabilities are known, the prospect is *risky,* and we write it as $x_p y$ with p the probability of x. Outcomes x and y are monetary gains, and x is always the better outcome. The *certainty equivalent* of a prospect is the sure outcome c equivalent to the prospect. The *matching probability* of an ambiguous prospect $x_E y$ is the probability m_E for which $x_E y$ and the risky prospect $x_{m_E} y$ are equivalent.

2.2. Non-Bayesian Choice Setup

We assume that two-outcome prospects $x_E y$ and $x_p y$ are evaluated as

$$\pi U(x) + (1-\pi) U(y), \tag{1}$$

where π is a *decision weight* that measures subjects' willingness to bet on the "winning event" and U is a strictly increasing *utility function*. As mentioned in the Introduction, for two-outcome prospects, Eq. 1 has the main models of decision under ambiguity and risk as special cases, including the ones that do best empirically (e.g., α-maxmin and prospect theory). For risk, $\pi = w(p)$, where w is a strictly increasing *probability weighting function* that maps probabilities onto the unit interval and satisfies $w(0) = 0$ and $w(1) = 1$. Under ambiguity, $\pi = W(E)$ where W is a *weighting function* that maps events to the unit interval $[0, 1]$ and satisfies $W(\emptyset) = 0$, $W(S) = 1$ and *monotonicity*: smaller sets (with respect to set inclusion) get smaller weights. In a Bayesian choice setup, i.e., under subjective expected utility (SEU), the willingness to bet on a given event E, $\pi = W(E)$, reduces to its subjective probability. This means that, in the standard Bayesian framework, rational decision makers can exhibit neither aversion nor proneness to ambiguity, i.e., ambiguity neutrality.

In our two studies, we consider different sources of uncertainty, including the risky source that was used as a baseline source. In Study 1, we characterize a source in terms of performance, be it a score or a rank, on a given test while keeping the difficulty of the test constant. In Study 2, we distinguish sources of uncertainty on two dimensions: performance (score or rank) and task difficulty (easy or hard). In both studies, we assume that subjects are probabilistically sophisticated within sources (Machina and Schmeidler 1992, Chew and Sagi 2006), but not between sources (which would amount to ambiguity neutrality). Probabilistic sophistication means that there exists an (additive) subjective probability measure $P(.)$ over the events. When $W(E) \neq P(E)$, we can find a strictly increasing *transformation function* f_s from $[0,1]$ to $[0,1]$ satisfying $f_s(0) = 0$ and $f_s(1) = 1$ such that

$$W(E) = f_s(P(E)). \tag{2}$$

The subscript s shows that the transformation function depends on the source of uncertainty. Note that, in the standard Bayesian choice setup, f_s should be the identity function, i.e., for all events E, $W(E) = P(E)$.

2.3. Willingness to Bet and Ambiguity Attitude

Abdellaoui et al. (2011) and Dimmock et al. (2016) define ambiguity aversion for a source s as the difference between f_s, the willingness to bet on source s, and w, the willingness to bet on chance. Dimmock et al. (2016) used matching probabilities to measure willingness to bet on events on the (objective) probability scale. In the Bayesian choice setup, which imposes ambiguity neutrality, the matching probability m_E of an event E should be equal to the subjective probability $P(E)$ of the event.

Under the more general model (1) with W as in equation (2), the equivalence between $x_E 0$ and $x_{m_E} 0$ gives

$$w(m_E) = f_s(P(E)),$$

which simplifies in

$$m_E = (w^{-1} \circ f_s)(P(E)). \tag{3}$$

Note that, in our non-Bayesian setup, this equation shows that m_E should be affected by biases related to overconfidence (through a possibly biased subjective probability $P(E)$) on the one hand and attitude towards ambiguity (through the difference between w and f_s at the likelihood level $P(E)$) on the other hand.

The comparison between a matching probability m_E and its corresponding subjective probability

$P(E)$ can be interpreted in terms of local ambiguity aversion, i.e., at the likelihood level $P(E)$. For any event E, if m_E is less than $P(E)$ then the subject is ambiguity averse. She is willing to give up $P(E) - m_E$ of her winning probability $P(E)$ to know the probability of winning and, thus, chance is more attractive than ambiguity. When m_E exceeds $P(E)$, the subject is ambiguity seeking. Thus, the function $w^{-1} \circ f_s$ reflects ambiguity attitudes. In the plot of m_E against $P(E)$, the *ambiguity function* $w^{-1} \circ f_s$ lies everywhere below the 45-degree for ambiguity averse subjects and above the 45-degree line for ambiguity seeking subjects.

We further split ambiguity attitudes into attitudinal pessimism (also referred to as ambiguity aversion) and likelihood insensitivity (Abdellaoui et al. 2011, Dimmock et al. 2016), which can be related to ambiguity perception (Baillon et al. 2018). More pessimistic attitudes means less weight to the best outcome. More likelihood insensitivity means less sensitivity to changes in likelihood.

To measure attitudinal pessimism and likelihood insensitivity, Dimmock et al. (2016, p. 1367) proposed the following simple model:

$$m_E = c + dP(E), 0 < P(E) < 1, \tag{4}$$

where the matching probability of an event E is linearly related to its subjective probability $P(E)$. Attitudinal pessimism and likelihood sensitivity in this model are measured by means of the indexes $b = 1 - d - 2c$ and $a = 1 - d$, respectively. Index b is inversely related to the average height of the regression line, hence representing a global index of ambiguity aversion. A smaller b means more optimistic attitudes. Index a reflects likelihood sensitivity with values smaller than 1, indicating a lack of sensitivity to probabilities. Figure B.1 (in Appendix B) illustrates the two indexes.

3. Elicitation Method

The present section describes our elicitation method. Specifically, it explains how beliefs, ambiguity attitudes, and overconfidence are measured in our two studies.

3.1. Beliefs

We measured subjective probabilities $P(E)$ using exchangeable events (Baillon 2008, Abdellaoui et al. 2011). Figure 1 illustrates the method. For any probability $p \in [0,1]$, let $s_p \in [0,100]$ denote the score such that a subject believed that the probability of obtaining at most score s_p was $p : P(\text{score} < s_p) = p$. By definition, $s_0 = 0$ and $s_1 = 100$. We first elicited the score $s_{0.5}$ that made subjects indifferent between betting on $100_{[s_0, s_{0.5})}0$ and betting on $100_{[s_{0.5}, s_1]}0$. By Eq. (2), $f_s(P([s_0, s_{0.5}))) = f_s(P([s_{0.5}, s_1]))$, i.e., equal willingness to bet on $[s_0, s_{0.5})$ and $[s_{0.5}, s_1]$. Because f_s is strictly increasing, we have $P([s_0, s_{0.5})) = P([s_{0.5}, s_1]) = 0.5$. Thus we obtain a clean measurement of subjective probabilities, free from ambiguity attitudes. The exchangeability method filters out the source-dependent distortion of subjective probabilities, which reflects attitudes towards ambiguity (cf. Eq. (2) in Section 2.2). Consequently, the score $s_{0.5}$ splits the whole score domain $[0, 100] = [s_0, s_1]$ into two equally likely subevents $[s_{0,}, s_{0.5})$ and $[s_{0.5}, s_1]$. Figure A.2 in Appendix A shows an example of a question we used to elicit $s_{0.5}$.

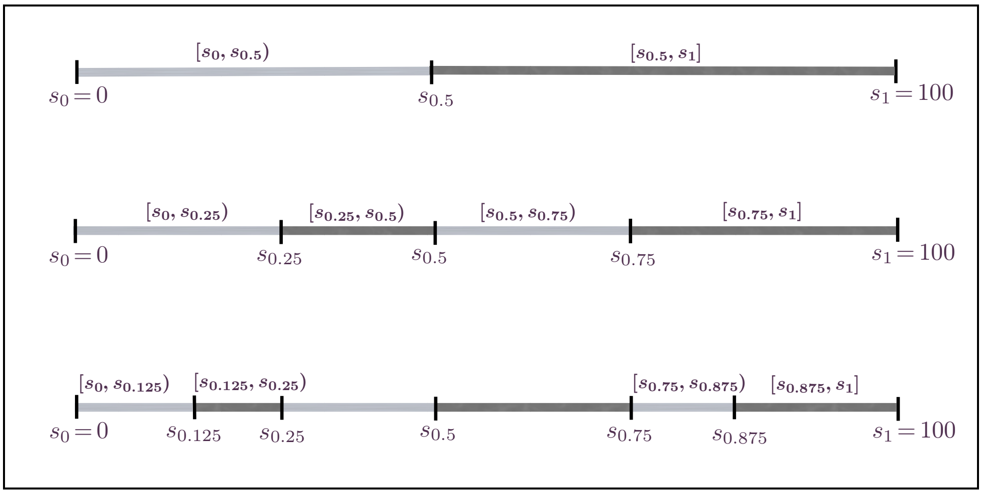
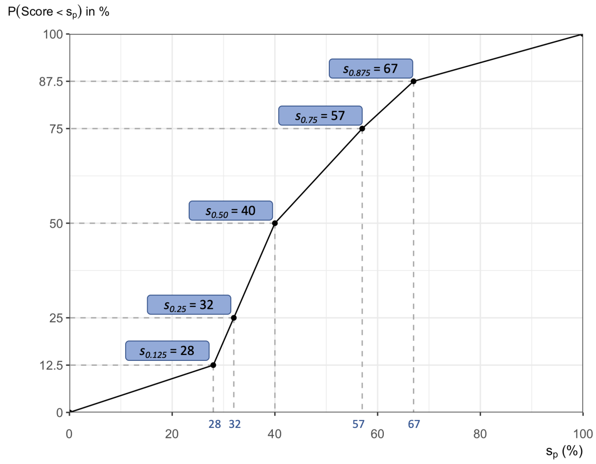


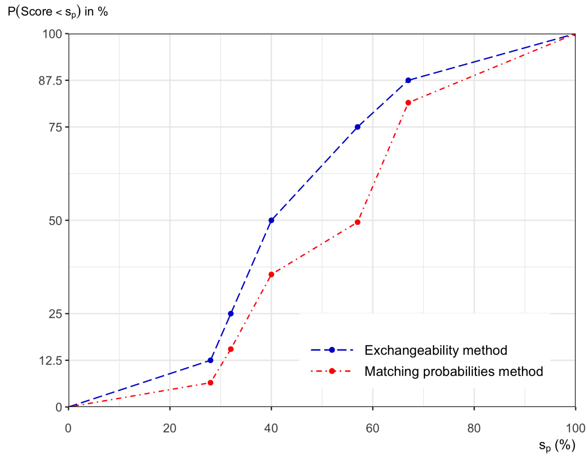
Figure 1: Illustration of the Exchangeable Events Method

We then elicited the score $s_{0.25}$ that splits the event $[s_0, s_{0.5})$ into two equally likely subevents by eliciting the indifference between $100_{[s_0, s_{0.25})}0$ and $100_{[s_{0.25}, s_{0.5})}0$, and the score $s_{0.75}$ that splits the event $[s_{0.5}, s_1]$ into two equally likely subevents. This led to four events $[s_0, s_{0.25})$, $[s_{0.25}, s_{0.5})$, $[s_{0.5}, s_{0.75})$ and $[s_{0.75}, s_1]$ with the same subjective probability 0.25. We also split the events $[s_0, s_{0.25})$ and $[s_{0.75}, s_1]$ into two equally likely subevents and so obtained 5 fractiles $s_{0.125}$, $s_{0.25}$, $s_{0.5}$, $s_{0.75}$, $s_{0.875}$ of the subject's subjective probability distribution. To illustrate, Figure 2 shows the distribution of subject 9 in Study 1. We used the same method to elicit 5 fractiles $r_{0.125}$, $r_{0.25}$, $r_{0.5}$, $r_{0.75}$, $r_{0.875}$ of subjects' subjective probability distributions about their rank. All fractiles were determined by a bisection process that zoomed in on the desired indifference value and determined it up to a precision of 1% for score and 1 point for rank (see Appendix A3, for an example).

To smooth out response errors, we also estimated the subjective probability distributions over the score and rank of each subject by a Beta distribution $B(\alpha, \beta)$. For estimation details see Appendix B1. We chose the Beta distribution for its flexibility (Berry 1996).



(a) Distribution of Beliefs (i.e. EE-based probabilities) about her Score



(b) Distribution of EE-based and MP-based Probabilities about her Score

Figure 2: Subject 9's Distribution of Subjective and Matching Probabilities about her Score

Note that the MP-based probabilities in Figure 2b correspond to the complementary events of the ones used in the elicitation of matching probabilities. The red curve being under the blue curve, therefore, suggests more optimistic beliefs with MP-based than EE-based probabilities.

3.2. Ambiguity Attitudes

For each elicited fractile s_{p_k}, we determined the matching probability m_k that led to indifference between the prospects $100_{m_k}0$ and $100_{[s_{p_k},s_1]}0$. In other words, subjects bet on their score being at least s_{p_k}. For example, for subject 9 in Study 1, we elicited the matching probabilities of $100_{[28,100]}0$, $100_{[32,100]}0$, $100_{[40,100]}0$, $100_{[57,100]}0$, and $100_{[67,100]}0$ where the events had subjective probabilities of 0.875, 0.75, 0.5, 0.25, and 0.125. Figure A.3 in Appendix A shows the elicitation of the matching probability of $100_{[33,100]}0$. We measured each matching probability using a bisection process with a precision of 1%.

Figure 2b shows the relation between the subjective probabilities and the matching probabilities for subject 9. As introduced in Section 2.3, the difference between the matching probability and the subjective probability reflects ambiguity attitude. For subject 9, the matching probabilities exceeded the subjective probabilities, and, consequently, this subject was ambiguity seeking.

To show the impact of ambiguity attitudes, we also estimated the beta distribution of the matching probabilities for each subject. This distribution is only equal to the subjective probability distribution when the subject is ambiguity neutral. Any difference between the distribution based on exchangeable events (*EE-based*) and matching probabilities (*MP-based*) is due to ambiguity attitudes. Much of the (behavioral economics) literature assumes a Bayesian setup where matching probabilities are assumed to equal to subjective probabilities despite the possibility that they could be affected by ambiguity attitudes in real choice situations. Figure 2b shows that this leads to a bias for subject 9. For example, while the estimate of the subject's probability of getting at least 40% of the questions right is 50% using exchangeable events, it is 62% using matching probabilities. Hence, for this subject, matching probabilities suggest more optimistic beliefs than with the exchangeable events.

3.3. Overconfident Beliefs and Overconfident Behavior

In our setup, overconfident behavior can be measured in two fashions. The first uses the probability scale and focuses on individual events, while the second uses the performance scale and resorts to the whole elicited probability distributions (score/rank).

At the level of a "performance-based" event E, we define overconfident behavior as the difference between willingness to bet on event E, m_E, and its true probability, $P_t(E)$. This difference represents a deviation from Bayesian behavior that can be decomposed as follows

$$m_E - P_t(E) = \underbrace{[P(E) - P_t(E)]}_{OB} + \underbrace{[m_E - P(E)]}_{AB}$$

where OB stands for the overconfidence bias and AB represents the ambiguity bias. In the absence of these two biases, i.e., $OB = AB = 0$, the observed matching probability of event E should be equal to its true probability, resulting in the absence of overconfident behavior. Note that, under the Bayesian choice setup, $AB = 0$.

Alternatively, overconfident behavior can also be defined in terms of the difference between expected performance (score/rank) as inferred from the elicited matching probability-based (MP-based) probability distribution over performance on the one hand and the observed actual performance on the other hand. Paralleling the above decomposition in terms of probabilities, this measure of overconfident behavior can further be decomposed into beliefs and attitudes components. The belief component is the measure of overconfidence in beliefs: the difference between the expected performance measured using the exchangeability method, deemed to factor out attitude considerations from beliefs, and the actual performance. The attitude component is the difference between the MP-based and EE-based expected performance.

4. Study 1

To address the key assumption resulting from our non-Bayesian choice framework that overconfident behavior is impacted by both belief-based overconfidence and ambiguity attitude, Study 1 investigates willingness to bet on performance-based events.

4.1. Experiment

Subjects and incentives Subjects were 58 students with diverse academic backgrounds.[6] Data were collected through personal interviews that lasted, on average, 75 minutes. Subjects received a

[6] Study 1 was performed with two additional subsamples of subjects devoted to another study on the impact of positive or negative feedback on beliefs and attitudes. Overall, we interviewed 187 subjects.

participation fee of €15. They also had a 10% chance of playing out their ability test questions for real and getting 20 cents per correct answer. Moreover, ten subjects were selected to play out one of their choices for real.[7] We selected the choice that was played out for real randomly. In the end, 19 subjects were paid according to their answers on the ability test and received an extra €5.20 on average. The ten subjects who played out one of their choices for real received an extra €76 on average. No subject was both paid for their score and got to play out one of their choices for real, even though theoretically this was possible.

Part 1: Ability test

- 50 questions taken from the Raven's matrices test
- Estimate the score and rank after 25 questions and at the end of the test

Part 2: Bets on performance and on chance

- 10-minute video that explained the type of questions and the real incentive system

The order of the three blocks was randomized

- Block 1: Bets on score on the ability test
  - Measure of the beliefs: elicitation of five events with probabilities 0.125, 0.25, 0.5, 0.75 and 0.875
  - Matching probabilities of five events with probabilities 0.125, 0.25, 0.5, 0.75 and 0.875
- Block 2: Bets on rank on the ability test
  - Measure of the beliefs: elicitation of five events with probabilities 0.125, 0.25, 0.5, 0.75 and 0.875
  - Matching probabilities of five events with probabilities 0.125, 0.25, 0.5, 0.75 and 0.875
- Block 3: Bets on risky lotteries
  - Certainty equivalents of ten risky prospects

Figure 3: Summary of Study 1

First part: Ability test Figure 3 summarizes the design of Study 1. We first asked 50 questions from Raven's matrices test (Raven et al. 2003), which measures reasoning ability. We chose the Raven's matrices as they have frequently been used in studies on overconfidence (e.g., Bruhin et al. 2018, Burks et al. 2013, Herz et al. 2014). In addition, there is a large variance in the questions' level of difficulty, which makes them well suited for our experiment. Before starting, subjects got written instructions and a simple example (see Appendix F). They had to answer each question within 50 seconds.

Subjects assessed their performance after the first 25 questions and at the end of the test.[8] We asked subjects to judge their score (percentage of correct answers) and their rank (among 100 randomly selected subjects).[9] Subjects used a scrollbar to indicate the most likely intervals in which their score and rank would lie (see Figure A.1 in Appendix). Each possible interval had a width of 8. For score, the intervals were 0%–8%, 2%–10%,..., 90%–98%, 92%–100%; for rank they were 1–9, 2–10,..., 91–99, 92–100. We used the midpoint of the elicited intervals as the subject's judged score and rank. We did not use real incentives for these judgments.

[7]Subjects knew this before the experiment started.

[8]This was done to study the effect of learning, which is not reported here.

[9]In our experimental setup, subjects compare themselves with a non-individuated target (Alicke et al. 1995).

A_k	x^k	Event E_k	y^k	$P(E_k)$
A_1	100	$[s_{0.875}, s_1]$ $[r_0, r_{0.125}]$	0	0.125
A_2	100	$[s_{0.75}, s_1]$ $[r_0, r_{0.25}]$	0	0.25
A_3	100	$[s_{0.5}, s_1]$ $[r_0, r_{0.5}]$	0	0.5
A_4	100	$[s_{0.25}, s_1]$ $[r_0, r_{0.75}]$	0	0.75
A_5	100	$[s_{0.125}, s_1]$ $[r_0, r_{0.875}]$	0	0.875

Table 1: Ambiguous Prospects in Study 1

Second part: Subjective probabilities and ambiguity attitudes The second part consisted of three sets of choices: the first was devoted to eliciting willingness to bet on chance through certainty equivalents. The collected certainty equivalents are not used in the present paper; they are treated as fillers in Study 1. The second and the third sets of choices focused on eliciting subjective probabilities and matching probabilities. The order of the three sets was random, but the order of the choices within sets was always the same.

Subjects first watched a 10-minute video about the questions they would get and the real incentive system (screenshots in Appendix F). They then got two sets of practice questions. The first set elicited (i) the certainty equivalent of the risky prospect $100_{0.4}0$, (ii) the matching probability of the ambiguous prospect $100_{[1,40]}0$ where $[1,40]$ is the event that the subject's rank is between 1 and 40, and (iii) the score s that led to indifference between the ambiguous prospects $100_{[20,s)}0$ and $100_{[s,80]}0$. The second set of practice questions assessed subjects' perceived minimum and maximum score (s_{min} and s_{max}) and rank (r_{min} and r_{max}) in the ability test. These were not used in the analyses.

4.2. Results

Beliefs about performance For each subject in Study 1, we elicited EE-based (using exchangeable events) and MP-based (using matching probabilities) subjective probability distributions for the score on the one hand and the rank on the other hand. This allowed us to infer individual expected performance through the expectations of the corresponding estimated Beta distributions. Further, each subject provided a guess about her actual score and rank on the ability test, i.e., judged score and judged rank. Figure 4 shows the corresponding empirical decumulative distributions, defined in terms of the events subjects bet on. For simplicity, and to facilitate comparison with Study 2, we reparametrized the ordinal rank as a percentile rank: rank #1 out of 100 corresponds to the percentile rank 100%, and rank #100 out of 100 corresponds to the percentile rank 0%. The left-hand panel of the figure shows for score $x\%$ the probability of getting at least that score. The right-hand panel shows the probability of being ranked higher than the lowest $x\%$.

At the aggregate level, as suggested by Figure 4, we could not reject the null hypothesis that subjects' MP-based expected scores ($M = 54.7$, $SD = 15.2$) were equal to their actual scores ($M = 55.4$, $SD = 17.5$), $t(57) = -0.352, \mathrm{p} = 0.727$. A similar conclusion holds for MP-based expected ranks ($M = 52.2$, $SD = 14.6$) as compared to actual ranks ($M = 52.9$, $SD = 34.3$), $t(57) = -0.163$, $\mathrm{p} = 0.871$. Binomial tests confirmed these results for score ($\mathrm{p} = 0.694$) and rank ($\mathrm{p} = 0.358$). Using the definition from Section 3.3, the difference between expected performance as inferred from the matching probabilities and the actual performance measures overconfident behavior. Therefore, we did not observe overconfident behavior for score or rank in the sample. This shows that, in Study 1, an analysis assuming a Bayesian choice framework (i.e., SEU) would conclude that overall beliefs were well-calibrated, neither over- nor underconfident.

In contrast to MP-based beliefs, the EE-based approach to beliefs (which does not postulate SEU) points to underconfidence. Specifically, our observations reject the null hypothesis that expected scores ($M = 50.1$, $SD = 15.5$) were equal to actual scores ($M = 55.4$, $SD = 17.5$), $t(57) = -2.499, \mathrm{p} = 0.015$. Further, 69% of subjects were underconfident for score (binomial test, $\mathrm{p} = 0.005$). As for rank, although the mean of expected ranks ($M = 48$, $SD = 15.5$) was below that of actual ranks ($M = 52.9$, $SD = 34.3$), our observations cannot reject the null hypothesis that they were equal, $t(57) = -1.213, \mathrm{p} = 0.230$. That said, most subjects (66%) had an expected rank below their actual rank, which points to underconfidence (binomial test, $\mathrm{p} = 0.025$).

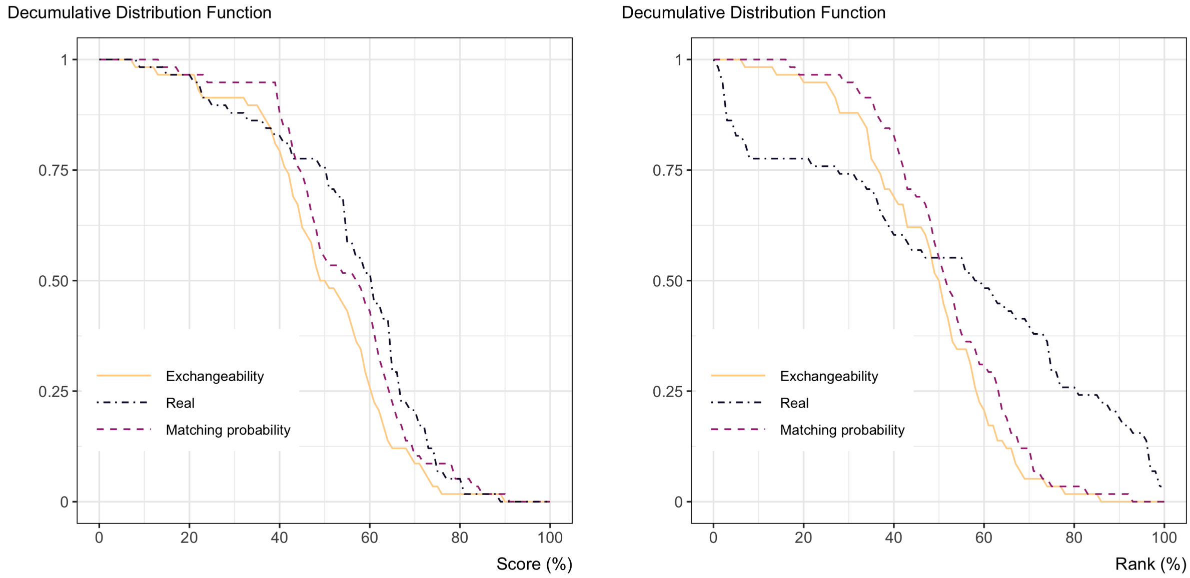
Figure 4 also suggests that EE-based expectations were lower than MP-based expectations for both the score and rank. As explained in Section 2, the observed gap can be explained in our non-Bayesian choice setup by ambiguity seeking. Paired t-tests reject the null hypothesis of equal expectations (EE-based vs. MP-based) for both score ($t(57) = -5.899$ $\mathrm{p} < 0.001$) and rank ($t(57) = -3.668$ $\mathrm{p} < 0.001$). Moreover, EE-based distributions were more precise than MP-based distributions in terms of standard deviation for both score ($t(57) = -3.464$ $\mathrm{p} = 0.001$) and rank, ($t(57) = -4.522$, $\mathrm{p} < 0.001$).[10]

[10]The mean of a Beta distribution $B(\alpha,\beta)$ is computed as $\alpha/(\alpha+\beta)$ and the standard deviation as $\sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2+(\alpha+\beta+1)}}$.

$P(E)$	m_E for score		m_E for rank	
	Mean (SD)	$\%(m_E > P(E))$	Mean (SD)	$\%(m_E > P(E))$
0.125	33.4 (17.6)	84.5%***	34 (19.7)	86.2%***
0.25	50.4 (21.6)	84.5%***	46-7 (19.1)	86.2%***
0.50	68.9 (18.7)	86.2%***	63.1 (16.7)	74.1%***
0.75	78.4 (16.3)	63.8%*	72.9 (18)	48.3%^{ns}
0.875	81.5 (16.7)	46.6%^{ns}	81.7 (14.8)	43.1%^{ns}

Binomial tests *** p<0.001, ** p<0.01, * p<0.05. Matching probabilities are expressed in percentages.

Table 2: Attitude Towards Ambiguity: m_E vs. $P(E)$



(a) Score

(b) Rank

Figure 4: Empirical Distributions of Real Performance and MP-based and EE-based Expectations assuming Beta Distributions

Ambiguity attitudes Given that MP-based probabilities exceeded EE-based probabilities, the dominant pattern was ambiguity seeking (Section 2.3). Table 2 shows that, on average, matching probabilities m_Es exceeded subjective probabilities $P(E)$s except for very likely events. At the individual level, we found ambiguity seeking in terms of score for subjective (i.e., EE-based) probabilities up to 0.75; in terms of rank, for $P(E) \leq 0.50$. We found more ambiguity seeking than studies using exogenous sources of uncertainty (e.g., Kocher et al. 2018). These studies typically also found ambiguity seeking for unlikely events but ambiguity aversion for moderately likely and likely events. This suggests that, within the so-called source effect, endogeneity may be considered as a factor that inflates optimistic attitudes. We investigate this issue in Study 2.

We further characterize ambiguity attitude in terms of attitudinal pessimism/optimism and insensitivity using the method described in Section 3, i.e., $m_E = c + dP(E)$ where $b = 1 - d - 2c$ and $a = 1 - d$ define the attitudinal pessimism and insensitivity indexes, respectively. Figure 5 displays the

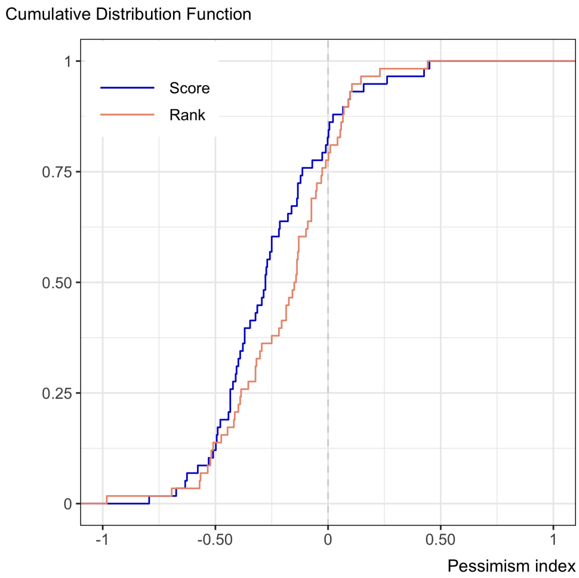


Figure 5: Empirical Distributions of the Attitudinal Pessimism Index

empirical Cumulative Distribution Functions of the attitudinal pessimism index for score and rank. It shows that 83% of subjects exhibited optimistic attitudes about their score (i.e., $b < 0$) and 78% about their rank (both binomial tests, $p < 0.001$). Additionally, although our observations point to marginally more optimistic attitudes about score than about rank, we could not reject the null hypotheses of equal attitudinal pessimism between score and rank ($t(57) = -1.978,\ p = 0.053$). The conclusion is similar and clearer for insensitivity ($t(57) = -0.317,\ p = 0.753$).

4.3. Conclusion of Study 1

Our first study demonstrates the feasibility and usefulness of unpacking overconfident behavior in a non-Bayesian choice framework. It shows, in particular, the role of ambiguity attitudes in explaining overconfident behavior. Overall, assuming SEU, we found neither under- nor overconfident behavior. This did not mean that subjects' beliefs were well-calibrated. Correcting for ambiguity attitudes, subjects were underconfident, particularly for their score. However, this underconfidence in beliefs was compensated by widespread ambiguity seeking for betting on one's performance, resulting in no under- or overconfident behavior.

The reason we found neither under- nor overconfident behavior in Study 1 may be that our design included both (very) easy and (very) hard questions and subjects found the questions overall neither hard nor easy. The empirical literature suggests that the degree of overconfidence depends on the difficulty of the task (Moore and Healy 2008): on easy tasks, people tend to underestimate their absolute performance but overestimate their relative performance, while on hard tasks, people tend to overestimate their absolute performance but underestimate their relative performance.

5. Study 2

This study aims to unpack overconfident behavior in a hard-easy context while assuming a non-Bayesian choice framework with real incentives. Among other things, this will allow us to check the robustness of the hard-easy effect in a non-Bayesian and incentivized setup. We also included an Ellsberg-like treatment to explore further the difference in ambiguity attitudes between endogenous and exogenous sources.

5.1. Experiment

Subjects and incentives We ran Study 2 online using Prolific.[11] Subjects received a participation fee of $5.50. In addition, they had the possibility of earning up to $7 extra based on their performance on the ability test or their choices in the experiment.

The experiment took on average 26 minutes. The design was similar to Study 1. We explain the differences below. Study 2 also had two parts: an ability test and a series of choices using subjects' performance on the ability task as events. Figure 6 summarizes the design of Study 2. Unlike Study 1, Study 2 further included an Ellsberg task to compare ambiguity attitudes for endogenous and exogenous sources. After having answered the ability test, subjects were told that one question from the second part of the study would be randomly selected to determine their bonus payment. Before each block of questions, we introduced the type of question and the procedure that would determine the bonus payment (see Appendix G). On average, participants earned a $3.3 bonus payment.

Part 1: Ability test

Between subject treatment: Easy vs. Hard test

- 20 questions taken from the Raven's matrices test

Part 2: Bets on performance and on chance

The order of the first two blocks was randomized

- Block 1: Bets on score on the ability test
  - Measure of the beliefs: elicitation of three events with probabilities 0.25, 0.5 and 0.75
  - Matching probabilities of three events with probabilities 0.25, 0.5 and 0.75 (randomized order)
- Block 2: Bets on rank on the ability test
  - Measure of the beliefs: elicitation of three events with probabilities 0.25, 0.5 and 0.75
  - Matching probabilities of three events with probabilities 0.25, 0.5 and 0.75 (randomized order)
- Block 3: Bets on the color of a ball drawn from an urn with unknown composition
  - Matching probabilities of three events with probabilities 0.25, 0.5 and 0.75 (randomized order)

End questions:

- Estimate the percentile score and percentile rank on the test
- Demographics questions

Figure 6: Summary of Study 2

[11] Due to COVID 19, we decided to run Study 2 online using our unpacking method with a simplified experimental protocol.

First part: Ability test The first part consisted of 20 questions from Raven's matrices test (Raven et al. 2003). Subjects had 45 seconds to answer each question. Subjects first answered a filter question. They could only proceed to the actual experiment if they answered this question correctly. Those who did not answer the question correctly received $1 for their participation. Before starting the experiment, subjects rated their perceived level of competence on the ability test on a 10-item Likert scale.

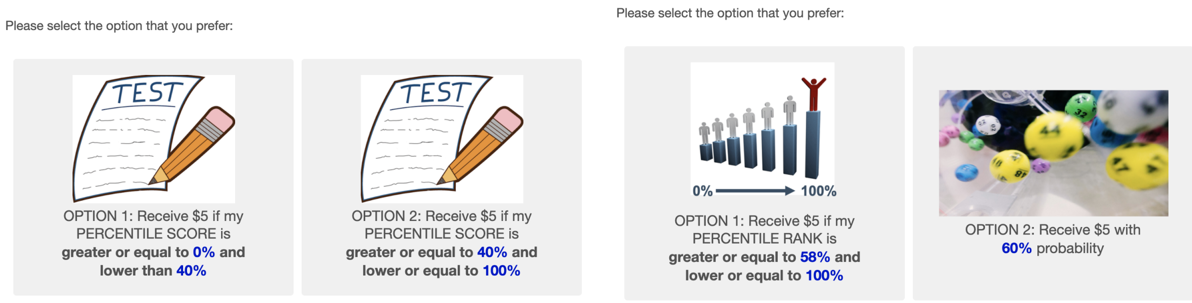
We randomly assigned subjects to one of two conditions: *easy* and *hard*. Subjects in the easy condition got 20 of the easiest questions from the Raven test; subjects in the hard condition got 20 of the hardest questions. We used information from Study 1 to design the two tests. To check the difficulty of the tests, we ran a pilot study with 45 subjects on Prolific. As we planned, the average actual score of the pilot was higher in the easy condition ($M = 87.5, SD = 15.3$) than in the hard condition ($M = 15.7, SD = 12.5$), $t(40.487) = 17.206,\ p < 0.001$.[12]

Second part: Subjective probabilities and ambiguity attitudes The second part consisted of three blocks of choices. The first two blocks were devoted to eliciting subjective probabilities and ambiguity attitudes for the score and rank, respectively. The order of these blocks was randomized, while the Ellsberg (third) block always came last. Within the three blocks, we randomized the order of the questions measuring ambiguity attitudes. The questions measuring subjective probabilities were not randomized due to the nature of the exchangeable events method. For all measurements, we used a bisection process with a precision of 5%. Each block started with an explanation of the task, a practice question (see Figure 7), and a series of comprehension questions included to check for the data quality (Appendix G).

We used the exchangeable events method to measure three fractiles of the subjective probability distribution for both score ($s_{0.25},\ s_{0.5},\ s_{0.75}$) and rank ($r_{0.25},\ r_{0.5},\ r_{0.75}$). For each of these fractiles, we measured the matching probabilities of getting at least that score or rank: $5_{[j_{0.25}, j_1]}0$, $5_{[j_{0.5}, j_1]}0$ and $5_{[j_{0.75}, j_1]}0$, $j = s, r$.

In the Ellsberg task, we elicited the matching probabilities of three ambiguous prospects using an urn with 100 balls, either blue, red, yellow, or green. The exact composition of the urn was unknown. Before the task, we showed four urns with different compositions to illustrate the concept of an urn with unknown composition. The events depended on the color of a ball randomly drawn from the urn. We elicited the matching probabilities of three events: $5_{[color \subset \{blue\}]}0$, $5_{[color \subset \{blue,\ red\}]}0$, and $5_{[color \subset \{blue,\ red,\ yellow\}]}0$. Like Dimmock et al. (2016), we assumed that subjects would consider each color equally likely a priori (see also Abdellaoui et al. 2011).

[12] All t-tests in Study 2 are two-sided. In case of unequal variance between the two samples, we report the Welch approximation to the degrees of freedom for two-sample t-tests.



(a) Beliefs for Score

(b) Willingness to Bet for Rank

Figure 7: Example of Questions used in Study 2

5.2. Results

Of the 408 subjects who participated in the experiment, 47 did not correctly answer the filter question and, hence, could not enter the experiment.[13] Consequently, 361 participants completed the experiment. Eighty percent of the subjects answered at least four of the five comprehension questions correctly. We ran the main analyses with these subjects. This left 155 participants in the easy condition and 134 in the hard condition. The results for the entire sample are similar and are in Appendix D.

Descriptive statistics Table D1 in the Appendix gives demographic statistics. Overall, there was no significant difference between the hard and easy groups in terms of gender, age, education, and prior perception of their competence in the ability test (all t-tests, $\text{p} > 0.1$). This suggests that the randomization worked. As expected, actual scores were higher on the easy ($M = 83.9,\ SD = 15.9$) than on the hard task ($M = 20.8,\ SD = 13.4$), $t(286.8) = 36.626,\ \text{p} < 0.001$.

Beliefs about performance As in Study 1, we computed, for each subject, the expected score and rank inferred from the subject's Beta distribution of subjective probabilities. We used EE-based probability distributions as they filter out attitudes toward risk and ambiguity.

Our observations show that EE-based expected performances were clearly affected by whether the subject was facing the hard or the easy condition. Two-sample t-tests showed that expected scores and ranks differed significantly between the two conditions ($t(287) = 14.727,\ \text{p} < 0.001$ for score and $t(287) = 11.023,\ \text{p} < 0.001$ for rank). Specifically, expected scores were higher in the easy condition ($M = 72.3,\ SD = 16.7$) than in the hard condition ($M = 42.1,\ SD = 18.1$). Similarly, expected ranks were higher in the easy condition ($M = 70.1,\ SD = 17.2$) than in the hard condition ($M = 47.3,\ SD = 18$).

[13] Participants were primarily informed that they could participate to the experiment only if they answered the filter questions.

In addition, two-sample t-tests showed that the precision of beliefs (as measured by standard deviation of the EE-based Beta distributions) was significantly different between the two conditions for both scores ($t(287) = -3.390,\ p < 0.001$) and ranks ($t(287) = -3.762,\ p < 0.001$). Subjects in the easy condition had more precise beliefs (i.e., smaller standard deviations) about their score ($M = 10,\ SD = 8.0$) than subjects in the hard condition ($M = 13.1,\ SD = 7.5$), as well as more precise beliefs about their rank ($M = 10.8,\ SD = 7.6$) than subjects in the hard condition ($M = 14.2,\ SD = 7.9$).

We also confirm the hard-easy effect in our non-Bayesian and incentivized choice setup. Paired t-tests rejected the null hypothesis of equality of expected and actual score in the easy condition ($t(154) = -8.59,\ p < 0.001$) and the analog hypothesis for rank ($t(154) = 4.52,\ p < 0.001$). The conclusion was similar in the hard condition for score and rank ($t(133) = 11.265,\ p < 0.001$ and $t(133) = -3.833,\ p < 0.001$, respectively). Specifically, in the easy condition, our subjects underestimated their expected score ($M = 72.3,\ SD = 16.7$) as compared to actual scores ($M = 83.9,\ SD = 15.9$), and overestimated their expected rank ($M = 70.1,\ SD = 17.2$) as compared to actual ranks ($M = 60.4,\ SD = 29.8$). For subjects in the hard condition, we observed the opposite pattern: expected ranks ($M = 47.3,\ SD = 18.0$) were lower than actual ranks ($M = 58,\ SD = 28.8$), and expected scores ($M = 42.1,\ SD = 18.1$) were higher than actual scores ($M = 20.8,\ SD = 13.4$).[14] The conclusions at the individual level of analysis accord with this overall picture. In the easy task, 78% of the subjects underestimated their score and 59% overplaced themselves (binomial test, $p < 0.001$ and $p = 0.024$, respectively). In contrast, in the hard condition, 82% overestimated their score (binomial test, $p < 0.001$) and 63% underplaced themselves (binomial test, $p = 0.002$).

In other words, when facing an easy test, subjects tended to overplace themselves relative to others and to underestimate their absolute performance. However, when facing a hard test, subjects tended to underplace themselves relative to others but to overestimate their absolute performance. These results point to the robustness of the hard-easy effect (Moore and Healy 2008) in a non-Bayesian setup. They also falsify the claim that the hard-easy effect could disappear when offering incentives (Murad et al. 2016, Grieco and Hogarth 2009). In addition, we observed that the difficulty of the test affected not only subjects' expected beliefs but also the precision of their beliefs. In Appendix D.2., we report similar results using the median of the beliefs (i.e., $s_{0.5}$) instead of the mean of the Beta distribution.

[14] The average percentile rank is greater than 50% as, following the literature, we did not break ties (e.g. Murphy and Weinhardt 2020). In addition, subjects who answered correctly the comprehension questions performed slightly better than those who failed some comprehension questions.

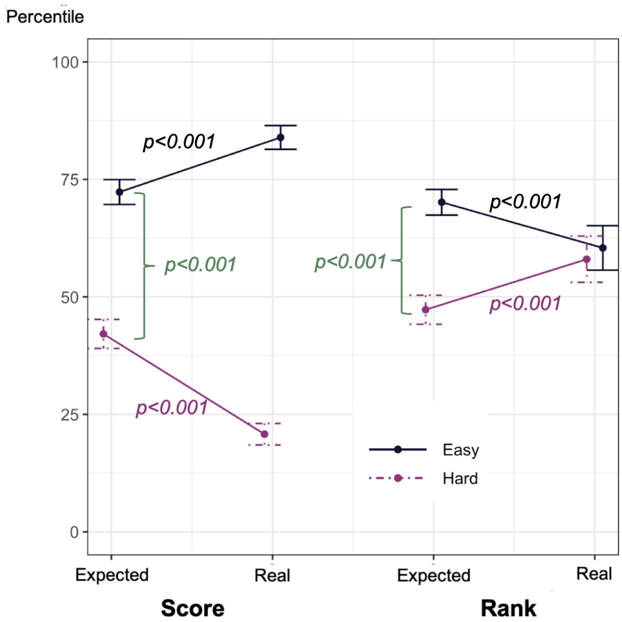


Figure 8: Expected (EE-based) vs. Actual Performance for Score and Rank

Ambiguity attitudes We analyzed ambiguity attitudes for different likelihood levels of the winning (performance-based) event E by comparing the matching probability m_E and the subjective probability $P(E)$. We first tested the null hypothesis of ambiguity neutrality—i.e., $m_E = P(E)$—in the two conditions using 2×3 ANOVA tests with repeated measures (matching vs. subjective probabilities $\times$ likelihood level). Specifically, our observations point to non-neutral ambiguity attitudes for both the score $(F(1,154) = 34.36$, $\text{p} < 0.001)$ and rank $(F(1,154) = 42.17$, $\text{p} < 0.001)$ in the easy condition, and for the score in the hard condition $(F(1,133) = 6.161, \text{p} = 0.014)$. Ambiguity neutrality turns out to be barely consistent with our observations for rank in the hard condition $(F(1,133) = 2.962, \text{p} = 0.088)$.

Table 3 reports the proportions of ambiguity seekers in the two conditions. For score in the easy condition, ambiguity seeking predominates when subjects face unlikely and moderately likely events, i.e., $P(E) \leq 0.5$. For the likely events, however, the shares of ambiguity seekers and ambiguity averters are similar. As for score in the hard condition, we observe a majority of ambiguity seekers for unlikely ($\text{p} < 0.001$) and moderately likely events ($\text{p} = 0.069$), and neither ambiguity seeking nor ambiguity aversion prevail for the likely events. For rank, the overall picture is similar in the easy condition, with ambiguity seeking prevailing for unlikely and moderately likely events, but we observe a different pattern in the hard condition. Specifically, the prevalence of ambiguity seeking for the unlikely event is accompanied by ambiguity aversion for the likely event.

Overall, this suggests that ambiguity attitude may depend on the task's difficulty when subjects

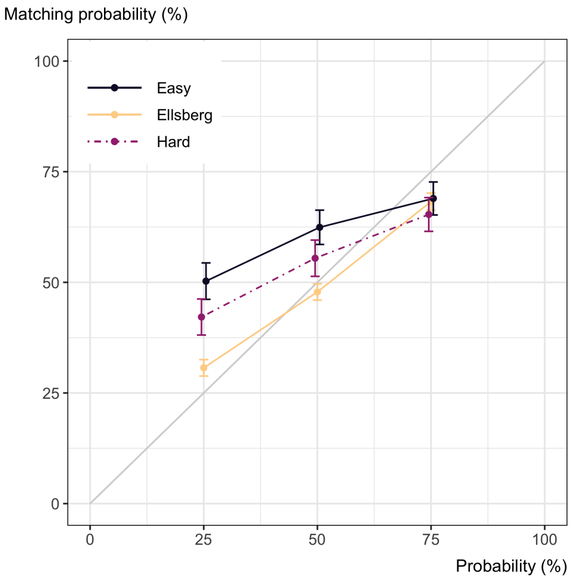
	$P(E)$	Easy		Hard	
		m_E	$\%(m_E > P(E)$	m_E	$\%(m_E > P(E)$
Score	0.25	50.27 (25.96)	81.9^{***}	42.16 (23.84)	76.1^{***}
	0.50	62.44 (24.43)	63.9^{***}	55.45 (23.97)	58.2^{+}
	0.75	68.95 (23.5)	49^{ns}	65.34 (22.33)	42.5^{ns}
Rank	0.25	51.31 (24.41)	80.6^{***}	42.46 (22.5)	76.1^{***}
	0.50	61.5 (23.97)	61.9^{**}	52.76 (22.68)	49.3
	0.75	69.66 (20.69)	47.7^{ns}	63.73 (24.16)	37.3^{**}

Binomial tests *** p<0.001, ** p<0.01, * p<0.05, $^{+}$ p<0.1. Matching probabilities are expressed in percentages.

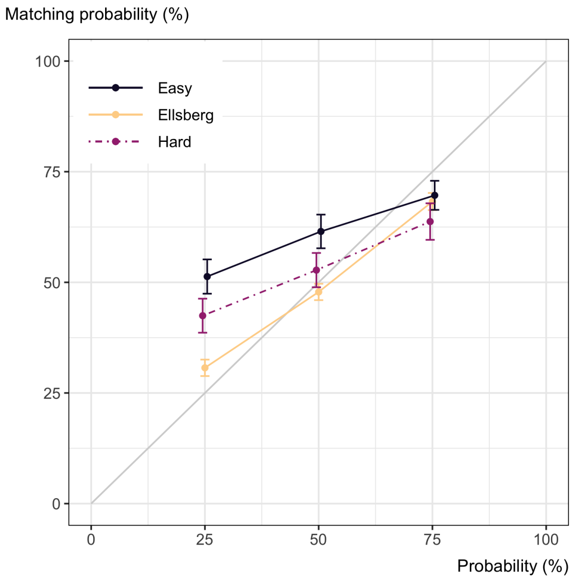
Table 3: Subjects Exhibiting Ambiguity seeking

compare their performance to others. In other words, this means that assuming a Bayesian choice framework would have resulted in ambiguity attitudes contributing to the hard-easy effect (at least for rank). To test for a possible effect of the task difficulty on ambiguity attitudes, we first ran a series of a 2×3 two-factor (condition $\times$ likelihood level) ANOVA tests. We could reject the null hypothesis that the matching probabilities for the score $(F(1,287) = 6.115,\ \mathrm{p} = 0.014)$ and the rank $(F(1,287) = 10.57,\ \mathrm{p} = 0.001)$ were the same under the two conditions. These results indicate more ambiguity seeking, for both score and rank, in the easy than in the hard condition. Post-hoc analyses show that the matching probabilities for the score were higher in the easy than in the hard condition for probability 0.25 $(F(1,287) = 7.56,\ \mathrm{p} = 0.018)$ and 0.5 $(F(1,287) = 5.98,\ \mathrm{p} = 0.045)$.[15] For the rank, the matching probabilities were higher in the easy than in the hard condition for probability 0.25 $(F(1,287) = 10.1,\ \mathrm{p} = 0.006)$, 0.5 $(F(1,287) = 10.0,\ \mathrm{p} = 0.006)$ and marginally higher for probability 0.75 $(F(1,287) = 5.05,\ \mathrm{p} = 0.075)$. In contrast, we did not observe differences in the matching probabilities for the Ellsberg urns $(F(1,287) = 2.170,\ \mathrm{p} = 0.142)$ between the two conditions.

[15] Post-hoc analyses were performed using Tukey's HSD tests with Bonferroni correction.



(a) MPs for Score



(b) MPs for Rank

Note: For Ellsberg curves, we grouped together the observations of the easy and hard conditions as there was no statistical difference between the two conditions.

Figure 9: Matching Probabilities for the Score and Rank

As explained earlier, following Dimmock et al. (2016), ambiguity attitudes can also be investigated in a parametric fashion through the regression of matching probabilities on their corresponding subjective probabilities, i.e., $m_E = c + dP(E)$. Figure 10 shows that most subjects had a negative attitudinal pessimism index $b = 1 - d - 2c$, which indicates ambiguity seeking. For the easy condition this proportion was significant both for score (67.1%, binomial: $p < 0.001$) and rank (63.2%, binomial: $p = 0.001$). For the hard condition, it was marginally significant for score (58.2%, binomial: $p = 0.069$) but not significant for rank (55.2%, binomial: $p = 0.261$). The graph for the easy condition lies everywhere above that of the hard condition, indicating more optimistic attitudes in the easy task than in the hard task. We could reject the null hypothesis that the attitudinal pessimism indexes were the same in both conditions for score, $t(287) = -2.473,\ p = 0.014$, and rank, $t(287) = -3.251,\ p = 0.001$. These results confirm a lower level of attitudinal pessimism in the easy condition than in the hard condition for both score ($M = -0.21,\ SD = 0.45$ vs. $M = -0.09,\ SD = 0.40$) and rank ($M = -0.22,\ SD = 0.41$ vs. $M = -0.06,\ SD = 0.40$).

We could also reject the null hypothesis that the insensitivity indexes for the score were the same in both conditions, $t(287) = 1.757,\ p = 0.080$ (see Figure D.1. in Appendix). The results suggested that insensitivity indexes for the score were marginally higher in the easy condition ($M = 0.63,\ SD = 0.43$) than in the hard condition ($M = 0.54,\ SD = 0.44$). Instead, we did not find any difference in the insensitivity toward rank between the two conditions, $t(287) = 1.158,\ p = 0.248$.

Condition	Source	Beliefs	Overall amb. attitudes	Likelihood-specific amb. attitudes		
				$P = 0.25$	$P = 0.50$	$P = 0.75$
Hard	Score	Overconfidence	Seeking^{+}	Seeking	Seeking^{+}	Neutrality
	Rank	Underconfidence	Neutrality	Seeking	Neutrality	Aversion
Easy	Score	Underconfidence	Seeking	Seeking	Seeking	Neutrality
	Rank	Overconfidence	Seeking	Seeking	Seeking	Neutrality

Note: The results for the overall attitudes are based on binomial tests on the attitudinal pessimism indexes.

$^{+}$: when the binomial test is only signifiant at 0.1.

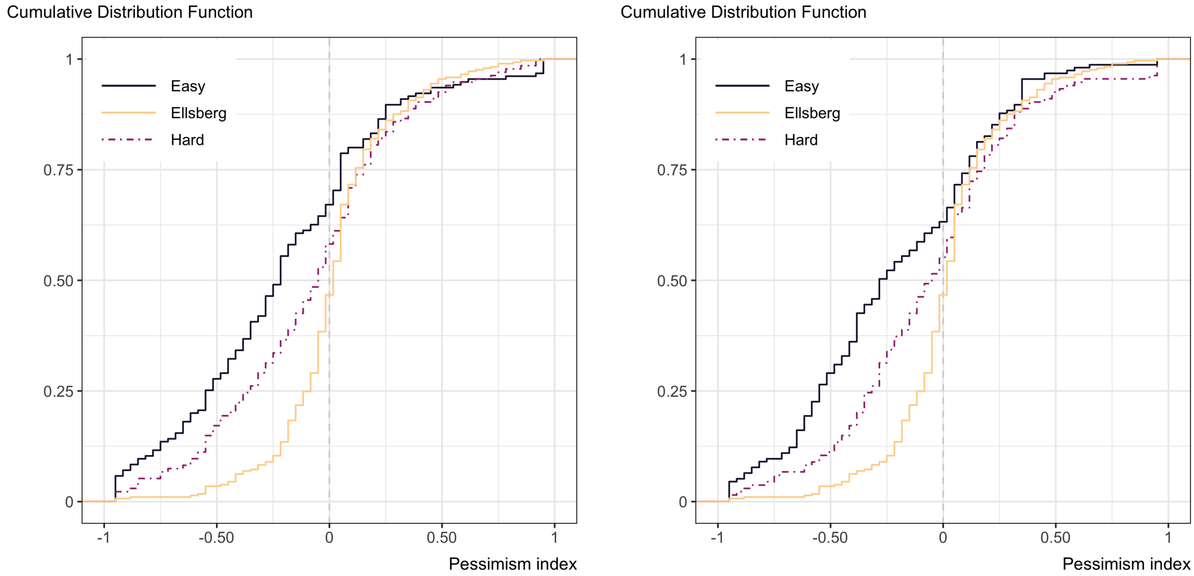
Table 4: Summary of the Study 2 Results: Beliefs and Ambiguity Attitudes

Ambiguity attitudes for endogenous and exogenous sources We also measured ambiguity attitudes for an exogenous source of uncertainty using an Ellsberg-like protocol. Figures 9 and 10 confirm our postulate in Study 1 of more ambiguity aversion for our exogenous source than for endogenous ones like score and rank. Our results for Ellsberg urns were similar to prior results in the literature. Using binomial tests, we found that the majority of subjects were ambiguity seeking at probability 0.25 (59.2%, $\mathrm{p} = 0.002$) and ambiguity averse at probabilities 0.5 (61.2%, $\mathrm{p} < 0.001$) and 0.75 (61.9%, $\mathrm{p} < 0.001$).

Using a series of paired t-tests, we can reject the null hypothesis of equal ambiguity aversion (as measured by the attitudinal pessimism index b) between the score and Ellsberg urns on the one hand, and the rank and Ellsberg urns on the other hand. Specifically, the null hypothesis was rejected for score ($t(154) = -7.196,\ \mathrm{p} < 0.001$) and for rank ($t(154) = -8.470, \mathrm{p} < 0.001$) in the easy condition. For the hard condition, the null hypothesis was rejected for score ($t(133) = -2.609,\ p = 0.010$) and was barely consistent with our observations for rank ($t(133) = -1.745,\ \mathrm{p} = 0.083$).

At the individual level, in terms of the attitudinal pessimism index, binomial tests show that the majority of the subjects in the easy condition exhibited more optimistic attitudes toward their score (78%, $\mathrm{p} < 0.001$) and rank (77%, $\mathrm{p} < 0.001$) than toward the Ellsberg's urns. Similarly, the majority of the subjects in the hard condition exhibited more optimistic attitudes toward their rank (60%, $\mathrm{p} = 0.030$) than toward the Ellsberg's urns. But only 57% of the subjects in the hard condition exhibited more optimistic attitudes toward their score than toward the Ellsberg's urns ($\mathrm{p} = 0.100$).

We also observed that subjects were less sensitive to the likelihood of the winning event when betting on score and rank than when betting on Ellsberg urns. This held for both the easy and the hard condition (paired t-test, all $p < 0.001$, see Table D.2 in Appendix). At the individual level, binomial tests show that the majority of the subjects were less sensitive to the likelihood of the winning event for their score (72%, $\mathrm{p} < 0.001$ in the easy condition and 73%, $\mathrm{p} < 0.001$ in the hard condition) and rank (67%, $\mathrm{p} < 0.001$ in the easy condition and 75%, $\mathrm{p} < 0.001$ in the hard condition) than toward the Ellsberg's urns. This suggests that the endogeneity of the source also affects likelihood insensitivity. Abdellaoui et al. (2011) observed no difference in likelihood insensitivity when comparing distinct exogenous sources of uncertainty and suggested that source preference shows up in the attitudinal pessimism index. This seems not true when comparing exogenous and endogenous sources.



(a) Attitudinal Pessimism - Score (b) Attitudinal Pessimism - Rank

Note: For Ellsberg curves, we grouped together the observations of the easy and hard conditions as there was no statistical difference between the two conditions.

Figure 10: Attitudinal Pessimism Indexes for Score and Rank

Decomposition of overconfident behavior: beliefs vs. attitude Finally, we studied how subjective probabilities and ambiguity attitudes combined in producing overconfident behavior, using the decomposition explained in Section 3. We computed for each subject the ratio between the ambiguity attitude bias (i.e., MP-based expectation minus EE-based expectation) and belief bias (i.e., EE-based expectation minus actual performance) parts to get insight into their relative contribution in producing overconfident behavior. In the easy condition, the median ratios were 33.5% (IQR=[15.6%; 70.5%]) for score and 28.2% (IQR=[13.1%; 66.3%]) for rank. In the hard condition, they were 22.0% (IQR=[7.8%; 69.3%]) for score and 19.1% (IQR=[6.0%, 44.3%]) for rank. These results suggest that, while the belief bias is larger than the ambiguity attitude bias, ambiguity attitudes play an important role in producing over- or underconfident behavior.

Competence effect Prior literature suggests that the perceived degree of competence can also affect willingness to bet on an uncertain event. People tend to exhibit more optimistic attitudes for more familiar sources (e.g., Kilka and Weber 2000, 2001) or when they feel competent about a source of uncertainty (Heath and Tversky 1991, Keppe and Weber 1995, Tversky and Fox 1995). To explore this mechanism, we ran analyses separating the subjects based on their *perceived competence* in the task. We classified as *high competence* subjects who rated their competence above the mean perceived competence in the sample (i.e., greater than 6.8) and the rest as *low competence*. Table D.3 (in Appendix) reports the results of a series of regression analyses of the attitudinal pessimism index on the experimental condition and the perceived competence on the task. We controlled for gender, age, and education in all the models.

Overall, the results confirm the effect of task difficulty on ambiguity attitudes. Controlling for perceived competence, subjects in the easy condition exhibited less pessimistic attitudes toward their score ($p = 0.013$, in model 1a) and rank ($p < 0.001$, in model 3a) than subjects in the hard condition.

Our analyses also confirm the results of the literature on the competence effect. Subjects who felt more competent on the task exhibited less pessimistic attitudes toward their rank ($p = 0.041$, in model 3a) than subjects who felt less competent. For the score, perceived competence was marginally related to lower pessimistic attitudes but only in the hard condition ($p = 0.090$, in model 1b). Finally, subjects who felt competent were less insensitive towards changes in likelihood, but only for their rank ($p = 0.010$).

Finally, we also analyzed the effect of participants' perceived competence on beliefs about performance (table D.4 in Appendix). Subjects who felt more competent on the task expected their score ($p < 0.001$, in model 1a) and rank ($p < 0.001$, in model 3a) to be higher than subjects who felt less competent. In contrast, we did not observe any effect of competence on the precision of beliefs.

6. Discussion

As stated by Kahneman (2011), "in terms of its consequences for decisions, [overconfidence] may well be the most significant of the cognitive biases" (p.255).[16] The importance of this behavioral trait is reflected in the large academic literature analyzing its effect in contexts as different as corporate decisions, entrepreneurship, and political behavior (Camerer and Lovallo 1999, Malmendier and Tate 2008, Ortoleva and Snowberg 2015). Assessing overconfidence is a difficult task as it requires measuring decision makers' beliefs. Both in the field and the lab, scholars have used methods based on revealed preferences to measure overconfidence. This paper reports experimental results showing that such measures may capture an attitude component on top of the belief component of overconfidence. Ignoring this attitude component may lead to mistaken conclusions about overconfidence in beliefs and, consequently, about the most adequate (policy) action to counter any overconfidence. To our knowledge, we provide the first investigation of overconfidence that elicits subjective probability distributions (in an EE-based fashion) for endogenous sources of uncertainty in a choice-based and incentivized fashion while circumventing ambiguity attitudes. In particular, this provides a robustness check of popular results based on stated beliefs (with or without scoring rules).

In two studies, we measured decision makers' beliefs and attitudes to the ambiguity surrounding their performance on an ability test. We used the participants' scores and ranks to assess overconfidence for absolute and relative performance, and we measured both stated and choice-based

[16]Note that Kahneman (2011) referred to overestimation, which he called optimism, and used overconfidence to refer to overprecision.

beliefs. We assumed a non-Bayesian choice framework, in the spirit of Abdellaoui et al. (2011) and Dimmock et al. (2016), where the willingness to bet on an event can differ from its subjective probability owing to, amongst other factors, non-neutral ambiguity attitudes.

In Study 1, all participants took a test composed of both hard and easy questions. This allowed us to test the method and assess the role of ambiguity attitudes on overconfident behavior. Study 2 was a between-subject experiment in which we manipulated the task's difficulty. This allowed us to investigate the hard-easy effect in a non-Bayesian and choice-based setup and, in particular, to explore whether a task's difficulty affected not only the level of overconfident beliefs but also ambiguity attitudes. Study 2 also measured ambiguity attitudes toward an exogenous source of uncertainty (i.e., Ellsberg urn). This has made it possible to test whether ambiguity attitudes for endogenous sources differed from those for exogenous sources commonly used in empirical studies. To better understand the effect of attitudes on overconfident behavior, we begin by discussing our findings related to ambiguity attitudes.

6.1. Attitudes when Betting on Oneself

We found more ambiguity seeking for our two endogenous sources of uncertainty (i.e., score and rank) than what has commonly been observed for exogenous sources of uncertainty. The existing literature typically finds ambiguity seeking for low levels of likelihood but (substantial) ambiguity aversion for intermediate and high levels of likelihood. Yet more than 75% of the subjects in Study 1 were ambiguity seeking for low and intermediate levels of likelihood both for score and rank. Around half of the subjects were still ambiguity seeking for high levels of likelihood for score.

In Study 2, we observed systematic optimistic attitudes for score for unlikely ($P = 0.25$) and moderately likely ($P = 0.50$) winning events and ambiguity neutrality for likely events ($P = 0.75$), independent of the difficulty of the test. For the rank, we also observed ambiguity-seeking behavior for unlikely events in both hard and easy tasks and for moderately likely events but only in the easy task. For likely events, ambiguity seeking gives way to a predominance of ambiguity neutrality and even ambiguity aversion for rank in the hard condition. Study 2 also allowed us to directly compare attitudes toward endogenous and exogenous sources of uncertainty. Most subjects exhibited more optimistic attitudes toward their score and rank than toward the exogenous source frequently used in the literature (i.e., Ellsberg's urns).

Derived while assuming a more general decision framework for both beliefs and attitudes, our results are consistent with the existing literature on endogenous sources of uncertainty, which has typically found evidence of ambiguity neutrality and ambiguity seeking. For instance, assuming SEU with a linear utility and without eliciting subjective probabilities, Keppe and Weber (1995) found patterns of ambiguity neutrality for uncertain events for which decision makers felt competent. Using stated subjective probabilities, Tversky and Fox (1995) found patterns consistent with ambiguity seeking as “participants preferred to bet on their uncertain beliefs in their area of competence rather than on known chance events.” (p. 281). Heath and Tversky (1991) also found similar patterns but acknowledged that their results could be due to “possible biases in the judgment process” (p. 20) related to non-additivity of probabilities estimates or non-neutral risk attitudes. Similar patterns of preference for ambiguity over risk have been observed in competitive settings, in which people bet on their performance relative to others (Klein et al. 2010, Gutierrez et al. 2020). Finally, the recent literature on ambiguity attitudes using simplified elicitation methods based on revealed preferences finds evidence of ambiguity seeking. Using three sources of uncertainty, Li, Turmunkh, and Wakker (2020) found evidence of ambiguity aversion for only one source. Their subjects exhibited less ambiguity aversion toward the uncertainty generated by other humans (choice of a partner in a game), be it strategic or not, than in an Ellsberg-like setup. Similarly, using stock price variation as a source of uncertainty, Baillon et al. (2018) found little ambiguity aversion at the aggregate level.

While many of the subjects in our two experiments exhibited ambiguity seeking for the different endogenous sources of uncertainty, we cannot generalize these patterns to other types of uncertainty. In our experimental settings, the uncertainty surrounding participants’ performance was mostly epistemic, i.e., the events were potentially knowable, but participants lacked knowledge concerning the true value of the event (Fox and Ülkümen 2011). People may behave differently when facing more aleatory uncertainty, in which the outcome is ex-ante unknowable. For instance, prior literature shows that people’s beliefs (Tannenbaum et al. 2017) and attitudes (Trautmann et al. 2008, Fox et al. 2021) depend on whether uncertainty is epistemic or aleatory. Therefore, we cannot generalize our finding to more aleatory types of uncertainty. We believe that studying how the two components of overconfident behavior are affected by the nature of uncertainty (epistemic vs. aleatory) is an exciting venue for future research.

6.2. Measuring Beliefs and Overconfidence using Choice-based Incentivized Tasks

Study 1 showed that committing to a Bayesian choice framework generated biased probabilistic beliefs about performance. Subjective probabilities inferred from matching probabilities under such a framework (Coutts 2019, Holt and Smith 2009, Urbig et al. 2009) exhibited discrepancies from those obtained through an elicitation method based on the exchangeability of events that filters out non-neutral attitudes towards ambiguity (Abdellaoui et al. 2011). Specifically, we found that measuring beliefs about performance under a Bayesian framework resulted in higher expected scores and ranks than under our model (using exchangeability). It also generated less precise beliefs (in terms of variance). Overall, this resulted in exaggerated belief-based optimism about performance. These apparent overconfident beliefs were caused by the positive effect of optimistic attitudes but were attributed to beliefs. Indeed, as seen earlier, we reported very optimistic attitudes toward the two endogenous sources of uncertainty: performance about score and rank.

These findings have important implications for the empirical literature on overconfidence. Indeed, scholars have increasingly used incentive-compatible measures of beliefs and, in particular, measures based on matching probabilities (Coutts 2019, Mobius et al. 2014, Urbig et al. 2009). While this method does not require elicitation of the utility function, it can lead to biased estimations of beliefs if attitude toward ambiguity is not taken into account. In particular, if people exhibit optimistic attitudes toward the source of uncertainty, as in our study, using similar methods along with the assumption of a Bayesian choice framework leads to overestimating overconfidence in beliefs. Some individuals would be classified as having overconfident beliefs when they are, in fact, either well-calibrated or underconfident.

Related to this, scholars in economics have suggested that using incentive-compatible measures to assess beliefs can lead to different empirical patterns than the ones obtained using stated measures (Blavatskyy 2009, Hoelzl and Rustichini 2005, Bruhin et al. 2018). For instance, Blavatskyy (2009) found that subjects exhibited underconfidence about their absolute performance when financial incentives were used. Similarly, Clark and Friesen (2009) reported a predominance of underconfidence (in particular for absolute performance) when "participants have incentives to forecast accurately." These studies use a variety of choice-based incentivized tasks to measure beliefs: scoring rules (Clark and Friesen 2009), eliciting certainty equivalents of bets (Murad et al. 2016), matching probabilities (Mobius et al. 2014, Bruhin et al. 2018), or voting games between lotteries and ambiguous events (Hoelzl and Rustichini 2005, Blavatskyy 2009, Grieco and Hogarth 2009).[17] One potential limitation of these methods is that participants' answers may be affected by non-neutral attitudes toward risk or the source of uncertainty generating performance (Hoelzl and Rustichini 2005). It is therefore important to further understand "how to assess and control the potential impact of ambiguity attitudes in the context of incentivized belief elicitation" (Murad et al. 2016, p. 39). Our paper offers a framework to measure beliefs while factoring out risk and ambiguity attitudes.

6.3. Extending the Hard-easy Effect Beyond Beliefs

Our paper also contributes to the extensive literature on the hard-easy effect. The degree of overestimation of one's absolute performance is affected by the difficulty of a task: people tend to underestimate their performance on easy tasks but overestimate it on hard tasks (Lichtenstein and Fischhoff 1977, Soll 1996). However, this pattern is reversed when looking at relative instead of absolute performance (see Table 4). Easy tasks tend to produce overplacement, and hard tasks tend to produce underplacement (Larrick et al. 2007, Moore and Small 2007, Kruger et al. 2008).

Unlike earlier studies that did not find support for the hard-easy effect using incentivized tasks (Murad et al. 2016, Grieco and Hogarth 2009), we found evidence of the hard-easy effect using our

[17] We note that the voting games between lotteries and ambiguous events are very similar to the elicitation of matching probabilities of ambiguous events.

non-Bayesian choice setup. Our method also allowed us to explore the hard-easy effect on the precision of the beliefs about performance. This is important as scholars have recently shown that measuring the distribution of beliefs instead of a point estimate extends our understanding of (over)confidence (Soll et al. 2021). Our results showed that task difficulty impacted both accuracy and precision of belief: easy tasks generated more precise beliefs about score and rank than hard tasks. Finally, we also found evidence of a hard-easy effect regarding ambiguity attitudes: the easy task generated more optimistic attitudes than the hard task. Although the difficulty of a task strongly affects beliefs, its effect on attitudes is not negligible.

As mentioned above, perceived competence has been identified as an element affecting attitudes toward ambiguity. People exhibit more optimistic attitudes when they feel competent about the source of uncertainty (Heath and Tversky 1991, Keppe and Weber 1995, Tversky and Fox 1995, Kilka and Weber 2001). In Study 2, we explored the competence effect by categorizing subjects based on their level of perceived competence on the task. Our results regarding the hard-easy effect on ambiguity attitudes are robust to controlling for perceived competence. Nevertheless, our analyses confirmed that subjects who felt more competent on the task were more ambiguity seeking for their rank than those who felt less competent. We found the same effect for the score but only in the hard task. Finally, we also found that perceived competence decreased the insensitivity but only for rank.

6.4. Conclusions and Future Directions

We proposed a method to separate the effects of beliefs and ambiguity attitudes on overconfident behavior. Our results show that assuming a non-Bayesian choice framework where beliefs are elicited without the interference of attitudes allows for an unpacking of overconfident behavior, including in a hard-easy setup. Specifically, we show that investigating overconfidence under a Bayesian choice framework could result in a distorted observation of overconfidence. Further, in addition to a choice-based confirmation of the hard-easy effect, we also show the existence of a hard-easy effect for ambiguity attitudes. We also show that attitude towards ambiguity is affected by whether the source of uncertainty is exogenous (Ellsberg) or endogenous (score and rank).

Assessing what causes overconfident behavior is important to determine the appropriate corrective actions. Interventions that can help correct beliefs may be less efficient for correcting attitudes. As an illustration, a growing literature in entrepreneurship proposes that entrepreneurship training can help reducing entrepreneurial biases, in particular overconfidence (Camuffo et al. 2020, Zhang et al. 2021). For these methods to be efficient on entrepreneurs' behavior, it is important to understand and quantify the main drivers of entrepreneurs' decisions, i.e., their beliefs and attitudes. If overconfident behavior is wrongly interpreted as overconfident beliefs (when in fact due to optimistic attitudes), correcting beliefs may not have the desired effect on behavior. If, instead, entrepreneurs have overconfident beliefs and at the same time strong pessimistic attitudes, it may not be necessary to debias their beliefs as biases in beliefs and attitudes may counterbalance each other. As an illustration, Study 2 shows that, on average, biases in beliefs and ambiguity attitudes reinforce each other for relative performance on easy tasks but counterbalance each other for absolute performance on easy tasks. In the latter case, it may be less efficient to attempt to correct beliefs. Finally, it is important to understand if/how the methods used to debias confidence in beliefs also affect attitudes. For example, if debiasing overconfidence in entrepreneurs' beliefs also reduces their attitude toward ambiguity, it may lead to undersupply or underinvestment.

The home bias—the fact that investors "invest disproportionately more in domestic stocks than standard portfolio theory would suggest as optimal allocation" (Lau et al. 2010, p. 191)—is another illustration of the importance of disentangling beliefs from attitudes. If the home bias is driven by biased beliefs (Kilka and Weber 2000, Solnik and Zuo 2017), providing information could help reduce this bias. If instead it is mainly driven by differences in attitudes toward uncertainty for local vs. distant investments (e.g., Chew et al. 2012), then methods designed to calibrate beliefs may not be effective. In this case, alternative methods such as using concrete language (Elliott et al. 2015) could be explored. Our paper cannot conclude on the particular questions of entrepreneurial biases or the home bias but highlights the importance of separating beliefs and attitudes and offers a method to do so.

Of course, our studies are not without limitations. We only focused on one type of task, i.e., logic puzzles measuring reasoning ability. While such tasks have been widely used in the overconfidence literature, future research could explore how both components of overconfident behavior are affected by the nature of the task. In our two studies, we mostly focused on epistemic sources of uncertainty: participants' absolute and relative performance on a test. For the relative performance, participants assessed their rank among a non-individuated target group with whom they had no personal contact. Prior literature shows that people's perception of uncertainty can depend on the nature of uncertainty, i.e., epistemic vs. aleatory (Tannenbaum et al. 2017) and, for relative performance, on whether the target group is individuated (Alicke et al. 1995). We believe our paper offers a method to explore how the different components of overconfident behavior affect decisions in settings involving aleatory uncertainty or comparison with an individuated target group. Finally, although we observed that the difficulty of a task affected not only individuals' beliefs but also their attitude towards the source of uncertainty, we did not identify the mechanisms responsible for this empirical pattern. A large literature has explored the mechanisms of the hard-easy effect on beliefs (e.g., Suantak et al. 1996). We believe that exploring the mechanisms that drive the hard-easy effect on ambiguity attitudes is a fruitful avenue for future research.

References

Abdellaoui, M., Baillon, A., Placido, L., and Wakker, P. P. (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review*, 101(2):695–723.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., and Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, 68(5):804.

Baillon, A. (2008). Eliciting subjective probabilities through exchangeable events: An advantage and a limitation. *Decision Analysis*, 5(2):76–87.

Baillon, A. and Bleichrodt, H. (2015). Testing ambiguity models through the measurement of probabilities for gains and losses. *American Economic Journal: Microeconomics*, 7(2):77–100.

Baillon, A., Huang, Z., Selim, A., and Wakker, P. P. (2018). Measuring ambiguity attitudes for all (natural) events. *Econometrica*, 86(5):1839–1858.

Berger, L., Emmerling, J., and Tavoni, M. (2017). Managing catastrophic climate risks under model uncertainty aversion. *Management Science*, 63(3):749–765.

Berry, D. A. (1996). *Statistics: a Bayesian perspective*. Number 04; QA279. 5, B4. Duxbury Press Belmont, CA.

Blavatskyy, P. R. (2009). Betting on own knowledge: Experimental test of overconfidence. *Journal of Risk and Uncertainty*, 38(1):39–49.

Bruhin, A., Santos-Pinto, L., and Staubli, D. (2018). How do beliefs about skill affect risky decisions? *Journal of Economic Behavior & Organization*, 150:350–371.

Bryan, K. A., Tilcsik, A., and Zhu, B. (2017). Which entrepreneurs are coachable and why? *American Economic Review*, 107(5):312–16.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. (2013). Overconfidence and social signalling. *Review of Economic Studies*, 80(3):949–983.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318.

Camuffo, A., Cordova, A., Gambardella, A., and Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science*, 66(2):564–586.

Chew, S. H., Ebstein, R. P., and Zhong, S. (2012). Ambiguity aversion and familiarity bias: Evidence from behavioral and gene association studies. *Journal of Risk and Uncertainty*, 44(1):1–18.

Chew, S. H. and Sagi, J. S. (2006). Event exchangeability: Probabilistic sophistication without continuity or monotonicity. *Econometrica*, 74(3):771–786.

Clark, J. and Friesen, L. (2009). Overconfidence in forecasts of own performance: An experimental study. *Economic Journal*, 119(534):229–251.

Coutts, A. (2019). Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics*, 22(2):369–395.

De Bondt, W. F. and Thaler, R. H. (1995). Financial decision-making in markets and firms: A behavioral perspective. *Handbooks in operations research and management science*, 9:385–410.

Dimmock, S. G., Kouwenberg, R., and Wakker, P. P. (2016). Ambiguity attitudes in a large representative sample. *Management Science*, 62(5):1363–1380.

Elliott, W. B., Rennekamp, K. M., and White, B. J. (2015). Does concrete language in disclosures increase willingness to invest? *Review of Accounting Studies*, 20(2):839–865.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, pages 643–669.

Feiler, D. and Tong, J. (2021). From noise to bias: Overconfidence in new product forecasting. *Management Science*.

Fox, C. R., Goedde-Menke, M., and Tannenbaum, D. (2021). Ambiguity aversion and epistemic uncertainty. *Available at SSRN 3922716*.

Fox, C. R. and Ülkümen, G. (2011). Distinguishing two dimensions of uncertainty. *in Essays in Judgment and Decision Making, Brun, W., Kirkebøen, G. and Montgomery, H., eds. Oslo: Universitetsforlaget.*

Galasso, A. and Simcoe, T. S. (2011). CEO overconfidence and innovation. *Management Science*, 57(8):1469–1484.

Gilboa, I. and Marinacci, M. (2016). Ambiguity and the bayesian paradigm. In *Readings in formal epistemology*, pages 385–439. Springer.

Gilboa, I., Postlewaite, A., and Schmeidler, D. (2011). Rationality of belief or: why savage's axioms are neither necessary nor sufficient for rationality. *Synthese*, 187(1):11–31.

Glaser, M., Langer, T., and Weber, M. (2013). True overconfidence in interval estimates: Evidence based on a new measure of miscalibration. *Journal of Behavioral Decision Making*, 26(5):405–417.

Grieco, D. and Hogarth, R. M. (2009). Overconfidence in absolute and relative performance: The regression hypothesis and bayesian updating. *Journal of Economic Psychology*, 30(5):756–771.

Gutierrez, C., Åstebro, T., and Obloj, T. (2020). The impact of overconfidence and ambiguity attitude on market entry. *Organization Science*, 31(2):308–329.

Heath, C. and Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1):5–28.

Herz, H., Schunk, D., and Zehnder, C. (2014). How do judgmental overconfidence and overoptimism shape innovative activity? *Games and Economic Behavior*, 83:1–23.

Ho, P.-H., Huang, C.-W., Lin, C.-Y., and Yen, J.-F. (2016). CEO overconfidence and financial crisis: Evidence from bank lending and leverage. *Journal of Financial Economics*, 120(1):194–209.

Hoelzl, E. and Rustichini, A. (2005). Overconfident: Do you put your money on it? *Economic Journal*, 115(503):305–318.

Holt, C. A. and Smith, A. M. (2009). An update on bayesian updating. *Journal of Economic Behavior & Organization*, 69(2):125–134.

Hooshangi, S. and Loewenstein, G. (2018). The impact of idea generation and potential appropriation on entrepreneurship: An experimental study. *Management Science*, 64(1):64–82.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Keppe, H.-J. and Weber, M. (1995). Judged knowledge and ambiguity aversion. *Theory and Decision*, 39(1):51–77.

Kilka, M. and Weber, M. (2000). Home bias in international stock return expectations. *The Journal of Psychology and Financial Markets*, 1(3-4):176–192.

Kilka, M. and Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty? *Management Science*, 47(12):1712–1726.

Klein, W. M., Cerully, J. L., Monin, M. M., and Moore, D. A. (2010). Ability, chance, and ambiguity aversion: Revisiting the competence hypothesis. *Judgment and Decision Making*, 5(3):192.

Kocher, M. G., Lahno, A. M., and Trautmann, S. T. (2018). Ambiguity aversion is not universal. *European Economic Review*, 101:268–283.

Kruger, J., Windschitl, P. D., Burrus, J., Fessel, F., and Chambers, J. R. (2008). The rational side of egocentrism in social comparisons. *Journal of Experimental Social Psychology*, 44(2):220–232.

Larrick, R. P. (2004). Debiasing. *Blackwell handbook of judgment and decision making*, pages 316–338.

Larrick, R. P., Burson, K. A., and Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes*, 102(1):76–94.

Lau, S. T., Ng, L., and Zhang, B. (2010). The world price of home bias. *Journal of Financial Economics*, 97(2):191–217.

Li, C., Turmunkh, U., and Wakker, P. P. (2020). Social and strategic ambiguity versus betrayal aversion. *Games and Economic Behavior*, 123:272–287.

Lichtenstein, S. and Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, 20(2):159–183.

Machina, M. J. and Schmeidler, D. (1992). A more robust definition of subjective probability.

Econometrica, pages 745–780.

Malmendier, U. and Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance*, 60(6):2661–2700.

Malmendier, U. and Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20–43.

McGraw, A. P., Mellers, B. A., and Ritov, I. (2004). The affective costs of overconfidence. *Journal of Behavioral Decision Making*, 17(4):281–295.

Mobius, M., Niederle, M., Niehaus, P., and Rosenblat, T. (2014). Managing self-confidence: Theory and experimental evidence. Technical report, UC San Diego.

Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502–517.

Moore, D. A. and Small, D. A. (2007). Error and bias in comparative judgment: on being both better and worse than we think we are. *Journal of Personality and Social Psychology*, 92(6):972.

Murad, Z., Sefton, M., and Starmer, C. (2016). How do risk attitudes affect measured confidence? *Journal of Risk and Uncertainty*, 52(1):21–46.

Murphy, R. and Weinhardt, F. (2020). Top of the class: The importance of ordinal rank. *The Review of Economic Studies*, 87(6):2777–2826.

Offerman, T., Sonnemans, J., Van de Kuilen, G., and Wakker, P. P. (2009). A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes. *The Review of Economic Studies*, 76(4):1461–1489.

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review*, 105(2):504–35.

Raiffa, H. (1961). Risk, ambiguity, and the savage axioms: Comment. *Quarterly Journal of Economics*, 75(4):690.

Raven, J., Raven, J. C., and Court, J. H. (2003). Manual for Raven's progressive matrices and vocabulary scales. Section 4, The advanced progressive matrices. *Harcourt Assessment*.

Soll, J. B. (1996). Determinants of overconfidence and miscalibration: The roles of random error and ecological structure. *Organizational Behavior and Human Decision Processes*, 65(2):117–137.

Soll, J. B., Milkman, K. L., and Payne, J. W. (2016). A user's guide to debiasing. *The Wiley Blackwell handbook of judgment and decision making*, 2:924–951.

Soll, J. B., Palley, A. B., and Rader, C. A. (2021). The bad thing about good advice: Understanding when and how advice exacerbates overconfidence. *Management Science*.

Solnik, B. and Zuo, L. (2017). Relative optimism and the home bias puzzle. *Review of Finance*, 21(5):2045–2074.

Suantak, L., Bolger, F., and Ferrell, W. R. (1996). The hard–easy effect in subjective probability calibration. *Organizational Behavior and Human Decision Processes*, 67(2):201–221.

Tannenbaum, D., Fox, C. R., and Ülkümen, G. (2017). Judgment extremity and accuracy under epistemic vs. aleatory uncertainty. *Management Science*, 63(2):497–518.

Trautmann, S. T., Vieider, F. M., and Wakker, P. P. (2008). Causes of ambiguity aversion: Known versus unknown preferences. *Journal of Risk and Uncertainty*, 36(3):225–243.

Tversky, A. and Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102(2):269.

Urbig, D., Stauf, J., and Weitzel, U. (2009). What is your level of overconfidence? a strictly incentive compatible measurement of absolute and relative overconfidence. *Discussion Paper Series/Tjalling C. Koopmans Research Institute*, 9(20).

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge university press.

Winkler, R. L. and Murphy, A. H. (1970). Nonlinear utility and the probability score. *Journal of Applied Meteorology and Climatology*, 9(1):143–148.

Zhang, S. X., Foo, M.-D., and Vassolo, R. S. (2021). The ramifications of effectuation on biases in entrepreneurship–evidence from a mixed-method approach. *Journal of Business Venturing Insights*, 15:e00238.